# Contents

## Protocols

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>An Individual Cognitive Stimulation Therapy App for People With Dementia and Their Carers: Protocol for a Feasibility Randomized Controlled Trial (e24628)</td>
<td>Harleen Rai, Justine Schneider, Martin Orrell</td>
<td>5</td>
</tr>
<tr>
<td>Virtual Reality Self-help Treatment for Aviophobia: Protocol for a Randomized Controlled Trial (e22008)</td>
<td>Jamie Fehribach, Marieke Toffolo, Ilja Cornelisz, Chris van Klaveren, Annemieke van Straten, Jean-Louis van Gelder, Tara Donker</td>
<td>16</td>
</tr>
<tr>
<td>De-Implementing Opioid Use and Implementing Optimal Pain Management Following Dental Extractions (DIODE): Protocol for a Cluster Randomized Trial (e24342)</td>
<td>DRindal, Stephen Asche, Jan Gryczynski, Sheryl Kane, Anjali Truitt, Tracy Shea, Jeanette Ziegenfuss, Robert Schwartz, Donald Worley, Shannon Mitchell</td>
<td>30</td>
</tr>
<tr>
<td>Effects of a Digital Mental Health Program on Perceived Stress in Adolescents Aged 13-17 Years: Protocol for a Randomized Controlled Trial (e25545)</td>
<td>Eliane Boucher, Haley Ward, Julia Stafford, Acacia Parks</td>
<td>39</td>
</tr>
<tr>
<td>A Web-Based Lifestyle Intervention Aimed at Improving Cognition in Patients With Cancer Returning to Work in an Outpatient Setting: Protocol for a Randomized Controlled Trial (e22670)</td>
<td>A Drijver, Jaap Reijneveld, Linda Wesselman, Martin Klein</td>
<td>55</td>
</tr>
<tr>
<td>User-Centered Counseling and Male Involvement in Contraceptive Decision Making: Protocol for a Randomized Controlled Trial (e24884)</td>
<td>Mahesh Karra, Kexin Zhang</td>
<td>68</td>
</tr>
<tr>
<td>A Gamified, Social Media–Inspired, Web-Based Personalized Normative Feedback Alcohol Intervention for Lesbian, Bisexual, and Queer-Identified Women: Protocol for a Hybrid Trial (e24647)</td>
<td>Sarah Boyle, Joseph LaBrie</td>
<td>102</td>
</tr>
<tr>
<td>A Technology-Based Physical Activity Intervention for Patients With Metastatic Breast Cancer (Fit2ThriveMB): Protocol for a Randomized Controlled Trial (e24254)</td>
<td>Siobhan Phillips, Payton Solk, Whitney Welch, Lisa Auster-Gussman, Marilyn Lu, Erin Cullather, Emily Torre, Madelyn Whitaker, Emily Izenman, Jennifer La, Jungwha Lee, Bonnie Spring, William Gradishar</td>
<td>117</td>
</tr>
</tbody>
</table>
A Live Video Program to Prevent Chronic Pain and Disability in At-Risk Adults With Acute Orthopedic Injuries (Toolkit for Optimal Recovery): Protocol for a Multisite Feasibility Study (e28155)
Ana-Maria Vranceanu, Jafar Bakhshaie, Mira Reichman, James Doorley, A Elwy, Cale Jacobs, Neal Chen, John Esposito, David Laverty, Paul Matuszewski, Amrreza Fatehi, Lucy Bowers, Mitchell Harris, David Ring. ................................. 136

Development of a Novel Intervention (Mindful Steps) to Promote Long-Term Walking Behavior in Chronic Cardiopulmonary Disease: Protocol for a Randomized Controlled Trial (e27826)
Daniel Litrownik, Elizabeth Gilliam, Peter Wayne, Caroline Richardson, Reema Kadri, Pamela Rist, Marilyn Moy, Gloria Yeh. ................................. 151

Uncovering the Bone-Muscle Interaction and Its Implications for the Health and Function of Older Adults (the Wellderly Project): Protocol for a Randomized Controlled Crossover Trial (e18777)
Cassandra Smith, Xuzhu Lin, David Scott, Tara Brennan-Speranza, Ahmed Al Saedi, Alba Moreno-Asso, Mary Woessner, Ebrahim Bani Hassan, Nir Eynon, Gustavo Duque, Itamar Levinger. ................................. 167

Electronic Health Record–Embedded Individualized Pain Plans for Emergency Department Treatment of Vaso-occlusive Episodes in Adults With Sickle Cell Disease: Protocol for a Preimplementation and Postimplementation Study (e24818)
Lingzi Luo, Allison King, Yvonne Carroll, Ana Baumann, Donald Brambilla, Christopher Carpenter, Joseph Colla, Robert Gibson, Siera Gollan, Greg Hall, Lisa Klesges, Abdullah Kutlar, Matthew Lyon, Cathy Melvin, Sarah Norell, Martina Mueller, Michael Potter, Rachel Richesson, Lynne Richardson, Gery Ryan, Lauren Siewny, Marsha Treadwell, Leslie Zun, Janelle Armstrong-Brown, Lisa Cox, Paula Tanabe. ................................. 182

The Geriatric Acute and Post-Acute Fall Prevention Intervention (GAPcare) II to Assess the Use of the Apple Watch in Older Emergency Department Patients With Falls: Protocol for a Mixed Methods Study (e24455)
Daniel Strauss, Natalie Davoodi, Margaret Healy, Christopher Metts, Roland Merchant, Swechya Banskota, Elizabeth Goldberg. ................................. 210

Development and Validation of Clinical Prediction Models for Surgical Success in Patients With Endometriosis: Protocol for a Mixed Methods Study (e20986)
Nadine Marlin, Carol Rivas, John Allotey, Julie Dodds, Andrew Horne, Elizabeth Ball. ................................. 221

An Intervention for the Transition From Pediatric or Adolescent to Adult-Oriented HIV Care: Protocol for the Development and Pilot Implementation of iTransition (e24565)
Amanda Tanner, Nadia Dowshen, Morgan Philbin, Kelly Rulison, Andres Camacho-Gonzalez, Susan Lee, Shamia Moore, J Fortenberry, Sophia Hussien. ................................. 230

Technology-Based Fall Risk Assessments for Older Adults in Low-Income Settings: Protocol for a Cross-sectional Study (e27381)
Ladda Thiamwong, Jeffrey Stout, Joon-Hyuk Park, Xin Yan. ................................. 240

Effects of Participating in a Research Project During the COVID-19 Pandemic on Medical Students’ Educational Routines and Mental Health: Protocol for a Web-Based Survey Study (e24617)

Generating Intervention Concepts for Reducing Adolescent Relationship Abuse Inequities Among Sexual and Gender Minority Youth: Protocol for a Web-Based, Longitudinal, Human-Centered Design Study (e26554)
Robert Coulter, Shannon Mitchell, Kelly Prangley, Seth Smallwood, Leyna Bonanno, Elizabeth Foster, Abby Wilson, Elizabeth Miller, Carla Chugani. ................................. 257
Pregnancy Outcomes and Child Development Effects of SARS-CoV-2 Infection (PROUDEST Trial): Protocol for a Multicenter, Prospective Cohort Study (e26477)
Geraldo Fernandes, Felipe Motta, Lizandra Sasaki, Angelo Silva, Andreza Miranda, Aleida Carvalho, Ana Gomides, Alexandre Soares, Agenor Santos Jr, Caroline Alves, Ciro Gomes, Clara Siracusa, David Araújo Jr, Dayde Mendonça-Silva, José Jesus, Karina Costa, Maria Castro, Patricia Kurizky, Paulo França, Rosana Tristão, Yacara Pereira, Luiz Castro, Alberto Zaconeta, Cleandro Albuquerque, Lícia Mota ........................................ 277

Digital Health Intervention to Increase Health Knowledge Related to Diseases of High Public Health Concern in Iringa, Tanzania: Protocol for a Mixed Methods Study (e25128)
Christine Holst, Felix Sukums, Bernard Ngowi, Lien Diep, Tewodros Kebede, Josef Noll, Andrea Winkler .................................................. 288

Autonomic Nervous System Maturation and Emotional Coordination in Interactions of Preterm and Full-Term Infants With Their Parents: Protocol for a Multimethod Study (e28089)
Christina Kourmara, Theano Kokkinaki, Giorgios Giannakakis, Katerina Koutra, Eleftheria Hatzidaki .................................................. 316

Pesticide Exposure of Residents Living Close to Agricultural Fields in the Netherlands: Protocol for an Observational Study (e27883)
Daniel Figueiredo, Esmeralda Krop, Jan Duyzer, Rianda Gerritsen-Ebben, Yvonne Gooijer, Henk Holterman, Anke Huss, Cor Jacobs, Carla Kivits, Roel Kruijne, Hans Mol, Arné Oerlemans, Pieter Sauer, Paul Scheepers, Jan van de Zande, Erik van den Berg, Marcel Wenneker, Roel Vermeulen .................................................. 328

Structure and Functioning of Acute Inpatient Psychiatric Units in Spain: Qualitative Study (e26214)
Roberto Rodríguez-Jiménez, Iluminada Corripio, Ricardo Campos, Mario Páramo, Manuel Franco-Martín, Estefanía Segura, Sergio González, José Martínez-Raga .................................................. 345

Epidemiology of Malaria in East Nusa Tenggara Province in Indonesia: Protocol for a Cross-sectional Study (e23545)
Robertus Guntur, Jonathan Kingsley, Fakir Islam .................................................. 352

Rehabilitation Needs, Service Provision, and Costs in the First Year Following Traumatic Injuries: Protocol for a Prospective Cohort Study (e25980)
Helene Soberg, Håkon Moksnes, Audry Anke, Olav Roise, Cecilie Røe, Eline Aas, Unni Sveen, Christine Gaarder, Pål Naess, Eirik Helseth, Hilde Dahl, Frank Becker, Marianne Levstad, Kristian Bartnes, Christoph Schäfer, Mari Rasmussen, Paul Perrin, Juan Lu, Torger Hellstrøm, Nada Andelic .................................................. 364

The EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) Study: Protocol for the Design, Implementation, and Evaluation of a Community-Based Combination HIV Prevention Intervention for Female Sex Workers in Baltimore, Maryland (e23412)
Bradley Silberzahn, Catherine Tomko, Emily Clouse, Katherine Haney, Sean Allen, Noya Galai, Katherine Footer, Susan Sherman .................................................. 374

Nutritional Implications of Baby-Led Weaning and Baby Food Pouches as Novel Methods of Infant Feeding: Protocol for an Observational Study (e29048)
Rachael Taylor, Cathryn Conlon, Kathryn Beck, Pamela von Hurst, Lisa Te Morenga, Lisa Daniels, Jill Haszard, Alison Meldrum, Neve McLean, Alice Cox, Lesieli Tukuatua, Maria Casale, Kimberly Brown, Emily Jones, Ioanna Katiforis, Madeleine Rowan, Jenny McArthur, Elizabeth Fleming, Ben Wheeler, Lisa Houghton, Aly Diana, Anne-Louise Heath .................................................. 390

Protocol for the Pregnancy During the COVID-19 Pandemic (PdP) Study: A Longitudinal Cohort Study of Mental Health Among Pregnant Canadians During the COVID-19 Pandemic and Developmental Outcomes in Their Children (e25407)
Gerald Giesbrecht, Mercedes Bagshawe, Melinda von Sloten, Anna MacKinnon, Ashley Dhillon, Marcel van de Wouw, Elnaz Vaghef-Mehrabany, Laura Rojas, Danielle Cattani, Catherine Lebel, Lianne Tomfohr-Madsen .................................................. 403

Development of an Emergency Department–Based Intervention to Expand Access to Medications for Opioid Use Disorder in a Medicaid Nonexpansion Setting: Protocol for Engagement and Community Collaboration (e18734)
Lauren Walter, Li Li, Joel Rodgers, Jennifer Hess, Rachel Skains, Matthew Delaney, James Booth, Erik Hess .................................................. 423
Original Papers

Digital Epidemiologic Research on Multilevel Risks for HIV Acquisition and Other Health Outcomes Among Transgender Women in Eastern and Southern United States: Protocol for an Online Cohort (e29152)
Andrea Wirtz, Erin Cooney, Megan Stevenson, Asa Radix, Tonia Poteat, Andrew Wawrzyniak, Christopher Cannon, Jason Schneider, J Haw, James Case, Keri Althoff, Elizabeth Humes, Kenneth Mayer, Chris Beyrer, Allan Rodriguez, Sari Reisner, American Cohort To Study HIV Acquisition Among Transgender Women (LITE) Study Group.

Subclinical and Clinical Outcomes in Patients Coinfected With HIV and Chronic Hepatitis B Virus From Clinical Outpatient Centers in France: Protocol for an Ambispective, Longitudinal Cohort Study (e24731)
Anders Boyd, Lorenza Dezanet, Raisha Kassime, Patrick Mialhes, Caroline Lascaux-Combe, Julie Chas, Pierre-Marie Girard, Joël Gozlan, Fabien Zoulim, Constance Delaugerre, Hayette Rougier, Karine Lacombe.
An Individual Cognitive Stimulation Therapy App for People With Dementia and Their Carers: Protocol for a Feasibility Randomized Controlled Trial

Harleen Kaur Rai¹, PhD; Justine Schneider², PhD, Prof Dr; Martin Orrell¹, PhD, Prof Dr

¹Division of Psychiatry and Applied Psychology, Institute of Mental Health, University of Nottingham, Nottingham, United Kingdom
²School of Sociology and Social Policy, Law and Social Sciences Building, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Harleen Kaur Rai, PhD
Division of Psychiatry and Applied Psychology
Institute of Mental Health
University of Nottingham
Jubilee Campus
Triumph Road
Nottingham, NG7 2TU
United Kingdom
Phone: 44 1157484252
Email: Harleen.Rai@nottingham.ac.uk

Abstract

Background: There is a need for more resources to support the cognition and quality of life of people with dementia. The individual cognitive stimulation therapy (iCST) app aims to provide cognitive stimulation and social interaction to people with dementia and carers through interactive touchscreen technology. The iCST app has been developed according to the principles of CST and iCST, which have previously shown to improve the cognition and quality of life of people with dementia and benefit the relationship between the person with dementia and his/her carer. The iCST app has also shown to improve the quality of the carer’s life.

Objective: The aim of this study is to evaluate the usability of the iCST app intervention and the feasibility of conducting a full-scale randomized controlled trial (RCT) to assess the clinical effectiveness of the iCST app intervention compared to that of treatment-as-usual for people with mild-to-moderate dementia.

Methods: We aim to recruit 60 people with mild-to-moderate dementia and their informal carers as dyads in a multi-center feasibility RCT with a treatment-as-usual control group. Both parties must be able to provide informed consent and participate in the intervention. Dyads will complete a baseline assessment that will include cognition and quality of life measures and they will subsequently be randomized (1:1) to the iCST app intervention in addition to usual care or to usual care only. All participants will be followed up at 5 weeks and at 11 weeks after the baseline assessments. A range of feasibility outcomes will be assessed, including recruitment and retention rates, intervention fidelity and usability, and acceptability of the outcome measures. A sample of the experimental group will be invited to a semistructured posttrial interview to further examine the experience of using the iCST app.

Results: This study received funding in May 2015 and obtained ethical approval in March 2018. Data collection began in November 2018 and was completed in March 2020 with a total of 61 dyads recruited. Data analyses are in progress and the final results are expected to be available in the spring of 2021.

Conclusions: This study will investigate whether it is feasible to conduct a full-scale RCT to evaluate the clinical effectiveness of the iCST app in comparison to that of usual care alone. In addition, this study will examine the usability of the iCST app. The data will provide information on potential modifications to be made to the intervention, study design, and study process.

Trial Registration: ClinicalTrials.gov NCT03282877; https://clinicaltrials.gov/ct2/show/NCT03282877
International Registered Report Identifier (IRRID): DERR1-10.2196/24628

(JMIR Res Protoc 2021;10(4):e24628) doi:10.2196/24628

https://www.researchprotocols.org/2021/4/e24628
KEYWORDS
dementia; cognitive stimulation therapy; touchscreen technology; feasibility trial; quality of life; mHealth; apps

Introduction

Background

Dementia poses a significant challenge to individuals in staying mentally stimulated and engaged. This is further exacerbated, given the lack of resources to support the cognition and quality of life (QoL) for people with dementia. Cognitive stimulation therapy (CST) is a nonpharmacological group treatment, which is strongly recommended by the National Institute for Health and Care Excellence. A previous randomized controlled trial (RCT) showed that CST can benefit the cognition and QoL of people with dementia [1]. The individual CST (iCST) intervention is delivered by a carer at home and has shown improvements in the relationship quality between the person with dementia and his/her carer and in the QoL of the carer [2].

Technology can improve accessibility to interventions by offering interventions on devices such as desktop computers and touchscreen tablets. For instance, in light of the COVID-19 pandemic wherein CST groups for people with dementia were unavailable, Cheung and Peri [3] found that it was feasible to offer virtual CST groups by using Zoom, a video conferencing software. This platform may support people with dementia to stay mentally stimulated and engaged in the safety of their homes. Touchscreen interventions, in particular, can have an array of benefits, as touchscreen platforms can be intuitive and therefore, have a high level of acceptance among people with dementia [4]. A systematic review has shown that touchscreen interventions with sound designs and tailored content can improve the well-being of people with dementia [5]. Moreover, these interventions can have a positive impact on the well-being of carers by decreasing their burden and improving the quality of the relationship with the person they are caring for by spending more time together. An example is the Computer Interactive Reminiscence and Conversation Aid (CIRCA) program, which offers computerized cognitive stimulation by prompting reminiscence among people with dementia supported by a wide range of multimedia stimuli on a touchscreen computer system. When used in a group setting, it can lead to improvements in both the cognition and QoL of people with dementia [6]. Researchers also found that CIRCA could positively benefit the relationship quality of the person with dementia and his/her carer when using the program together [7]. However, despite these encouraging findings, the field of computerized cognitive stimulation is underdeveloped [8], and high-quality studies involving carer-delivered computerized cognitive stimulation programs, in particular, remain scarce.

More research in this field is warranted as the existing evidence seems to be promising; there are minimal risks in engaging with computerized cognitive stimulation and such interventions could potentially be cost-effective. Therefore, researchers have developed a novel, touchscreen iCST app, which aims to offer cognitive stimulation to people with dementia and can be used at home with carers. Rather than using a touchscreen computer system like CIRCA, the iCST app has been developed as a native app; thus, it can be downloaded from the App Store and Google Play stores on touchscreen tablets, which further improves the accessibility of the intervention. Like CIRCA, the iCST app aims to stimulate conversation between the person with dementia and his/her carer by using a range of topics, which can include reminiscence. However, the iCST app also offers other types of activities supported by multimedia stimuli in order to target specific cognitive functions such as memory and language. It is hoped that the iCST app may produce combined benefits of engaging in CST, iCST, and touchscreen technology. The app’s development followed the Medical Research Council Framework for developing complex interventions and the Centre for eHealth Research Roadmap [9,10]. This included extensive reviews of CST and iCST materials and end-user involvement through informal consultations, focus groups, individual interviews, and usability questionnaires [11]. The next stage of development encompasses this feasibility RCT. Findings from this study will inform whether a large-scale RCT evaluating the clinical effectiveness of the iCST app is indicated by investigating relevant study parameters related to the study design and process.

Study Aim

The aim of this study is to evaluate the usability of the iCST app intervention and the feasibility of conducting a full-scale RCT to assess the clinical effectiveness of the iCST app intervention compared to that of treatment-as-usual (TAU) for people with mild-to-moderate dementia.

Methods

Study Design

The 26-item CONSORT checklist of information to include when reporting feasibility trials will be used for this study to ensure that all the necessary and relevant information is reported [12]. This study was registered with ClinicalTrials.gov on July 19, 2017 (registration number: NCT03282877). This study is a multi-center, feasibility RCT with an allocation ratio of 1:1. People with dementia and carers will be recruited as dyads and will be randomized to either the experimental (completing 30-minute iCST app sessions twice or thrice per week) or the TAU control group for 11 weeks. Dyads will complete the baseline assessment prior to randomization and thereafter, the first follow-up will be completed at 5 weeks after the baseline assessments and the second follow-up at 11 weeks after the baseline assessments (Figure 1). A sample of the experimental group will be invited for semistructured posttrial interviews to gain insights into the acceptability of the iCST app, including the experience of using the app and any facilitators and barriers for implementation in daily life.
Participants
Recruitment started on November 1, 2018 and is conducted in 5 secondary care settings in England: Derbyshire Healthcare National Health Service (NHS) Foundation Trust, Leicestershire Partnership NHS Trust, Lincolnshire Partnership NHS Foundation Trust, Northamptonshire Healthcare NHS Foundation Trust, and Nottinghamshire Healthcare NHS Foundation Trust. In addition, participants will be identified through a variety of settings, including general practitioner practices, community mental health teams, memory clinics, care homes, memory cafes, support groups, and voluntary sector organizations such as the Alzheimer’s Society. Remote recruitment will include registration on the website Join Dementia Research, publicizing the study using social media platforms such as Twitter, and the distribution of information leaflets and posters to organizations and professionals involved in the identification of possible participants.

Eligibility Criteria
The sample will include people with mild-to-moderate dementia and their informal carers (relatives or friends). The inclusion and exclusion criteria were adapted from previous CST and iCST research studies [1,2] and are shown in Textbox 1.
Inclusion and exclusion criteria for people with dementia and for their carers in this study.

**Inclusion criteria**

**For people with dementia**
- Meet Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria for dementia [13].
- Score 10 or above on the Mini-Mental State Examination [14] or score of 16 or above on the Montreal Cognitive Assessment [15], where available.
- Some ability to communicate and understand (eg, ability to give informed consent).
- Ability to speak and understand English.
- See/hear well enough to participate.
- No major physical illness or disability affecting their participation.
- Age range: 50 years–no maximum age limit.
- Availability of a touchscreen tablet for the person with dementia and for the carer.
- Availability of a carer (or relative/friend) to participate in the activities.

**For carers**
- Minimum age: 21 years.
- Ability to speak and understand English.
- See/hear well enough to participate.
- No major physical illness or disability affecting their participation.

**Exclusion criterion for people with dementia and for carers**

Concurrent participation in any other interventional study.

---

**Sample Size**

A formal sample size calculation is not appropriate for a feasibility trial. A previous audit of trials registered in the Clinical Research Network database in the United Kingdom found that most feasibility and pilot trials had a median of 30 or 36 participants per arm and the researchers recommend an upper limit of 60 participants for a feasibility trial [16]. Therefore, a target of 60 dyads is set for this study, leading to 30 dyads per treatment arm.

**Procedure**

**Screening for Eligibility**

Anticipating logistical support from the National Institute for Health Research Clinical Research Network East Midlands, we expect that staff members at each research site will check the eligibility of referrals received from clinicians and staff at the recruitment sources. Participants fulfilling the inclusion criteria will be sent a participant information sheet containing full details about the study. If the dyad is interested in participating, they will be recruited into the trial and a date for the baseline assessment and consenting will be set by the Clinical Research Network staff member.

**Randomization**

Randomization will take place after consent and the baseline assessments by using a web-based central randomization service called Sealed Envelope [17]. Block randomization will be employed with block sizes of 4 to 6 (randomly varied and generated by Sealed Envelope), which is a useful method for small sample sizes to allocate an equal number of participants to each treatment arm [18]. The researcher at the local site will perform each randomization using the participant identification code of the person with dementia only. The allocation to the experimental or TAU control group will automatically apply to the carer as well. Dyads will be informed of their allocation outcome over the telephone and, if necessary, a visit will be arranged for dyads in the experimental group to install the iCST app.

**Blinding**

This trial will include both blinded and unblinded researchers at each local site. It is not possible to blind the participants to their treatment arm as the iCST app is a nonpharmacological intervention. However, each study site will include at least one researcher kept in ignorance of study allocations. The baseline assessment can be performed by either researcher. However, both follow-ups will be completed only by the researcher who is unaware of the randomization outcome for each dyad. If disclosure does occur, this will be recorded by the visiting researcher along with details on how it occurred. The unblinded researcher will perform the randomization, communicate the outcome with the participants, and for the experimental group, install the iCST app, provide weekly telephone support calls, and complete the usability and acceptability questionnaire at the end of the study. Furthermore, the unblinded researcher will not be informed about the results of the assessments.

**iCST App Intervention**

Participants in the experimental group will use the iCST app (prototype v3.0) over 11 weeks after the baseline assessments. The content of the app was modified from the paper-based iCST manual, including the principles, themes, and activities [2], and
was based on consultations and qualitative research with people with dementia, carers, and the software development company [11]. The iCST app is a one-to-one, carer-led, home-based program of structured cognitive stimulation for people with dementia but delivered on a touchscreen tablet. It includes 21 activities consisting of both game-like interactive features such as audio-visual stimuli and discussion questions (Table 1). These activities offer mental stimulation not only through the content on the app but also through conversation.

Table 1. List of the individual cognitive stimulation therapy app activities (prototype v3.0).

<table>
<thead>
<tr>
<th>Type of activity</th>
<th>Activity name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Games only</td>
<td>• Being Creative</td>
</tr>
<tr>
<td></td>
<td>• Spaceman</td>
</tr>
<tr>
<td></td>
<td>• Trivia Quiz</td>
</tr>
<tr>
<td></td>
<td>• Word Search</td>
</tr>
<tr>
<td></td>
<td>• Sudoku</td>
</tr>
<tr>
<td></td>
<td>• Being Active</td>
</tr>
<tr>
<td></td>
<td>• Brainstorm</td>
</tr>
<tr>
<td>Discussion questions only</td>
<td>• Past Events</td>
</tr>
<tr>
<td></td>
<td>• Useful Tips</td>
</tr>
<tr>
<td></td>
<td>• My Life</td>
</tr>
<tr>
<td></td>
<td>• Arts</td>
</tr>
<tr>
<td></td>
<td>• Old Wives’ Tales</td>
</tr>
<tr>
<td></td>
<td>• Toys Are Us</td>
</tr>
<tr>
<td>Games and discussion questions</td>
<td>• Sounds</td>
</tr>
<tr>
<td></td>
<td>• Odd one Out</td>
</tr>
<tr>
<td></td>
<td>• The Price is Right</td>
</tr>
<tr>
<td></td>
<td>• Globe Trotter</td>
</tr>
<tr>
<td></td>
<td>• Food</td>
</tr>
<tr>
<td></td>
<td>• In Pairs</td>
</tr>
<tr>
<td></td>
<td>• Sayings</td>
</tr>
<tr>
<td></td>
<td>• I Spy</td>
</tr>
</tbody>
</table>

In addition, this app includes several other features such as a short introduction section explaining the background and key tips for using the app, a home screen that features completed activities, and a choice of 2 levels. Level 2 contains either more challenging content or different questions from Level 1, and it is up to the participants to determine which level they feel more comfortable with for each activity. Figure 2 contains screenshots of the iCST app. Considering previous CST and iCST research and findings from the development work, it is recommended that participants use the app for 2 or 3 times a week for 30 minutes [2]. Participants are free to spend more time on the app if they wish, and this will be recorded during the weekly telephone calls.
Training and Adherence

In order to ensure treatment integrity for all participants, individual study sites will receive a demonstration of the iCST app along with training in its installation and use prior to the start of recruitment. Dyads in the experimental group will receive an in-home visit from the unblinded researcher who will install the app and explain how it works using a short supplementary document containing instructions with screenshots of the app. Furthermore, all dyads will receive weekly telephone support calls from the same unblinded researcher in order to monitor adherence and to track the overall progress and any challenges or technical difficulties with using the iCST app. Phone calls will be completed with the carer and all questions will be included on a standardized telephone sheet. The questions relate to general experience, number of sessions completed in a week (on average), amount of time spent per session (on average), enjoyment, and any likes/dislikes. Any reason for not being able to use the iCST app over the week will also be recorded on the telephone sheet.

TAU Group

The control group will consist of the TAU group and will not receive any additional interventions. TAU groups are typically used to compare experimental interventions to care, which participants already receive in practice [19]. Therefore, the TAU control group will enable us to compare the effects of the iCST app with the natural progression of people with dementia under conditions of usual care. The treatments and services that are already available to people with dementia and their carers randomized to the TAU control group may differ between and within recruitment sites, for instance, regarding acetylcholinesterase inhibitors or being involved in some form of cognitive stimulation already. However, it is unlikely that people have access to computerized versions of iCST since these do not exist, to the best of our knowledge. The visiting researcher will record any current participation with CST groups and use of acetylcholinesterase inhibitors at the baseline assessment.

Outcomes

Feasibility Outcomes

In order to determine the feasibility of conducting a full-scale RCT with the iCST app in the future, this trial will investigate key feasibility aspects, including the rates of recruitment, screening, randomization, and retention by using enrolment logs [9]. Acceptability of the outcome measures will be evaluated by assessing the completion rates, and the acceptability and fidelity of the iCST app will be evaluated through weekly telephone support calls, analytics, and a usability and acceptability questionnaire. It is expected that >75% of the participants in the experimental group will need to complete the recommended minimum of 2 activities on average every week for the iCST app to be considered feasible. This benchmark has been adopted based on work in some previous feasibility trials, including psychological treatments where benchmarks for successful adherence ranged between 75% and 80% [20,21].
Clinical Outcomes

Previous CST and iCST research and the Interdem consensus statement on outcome measures for dementia informed the outcome measure selection for this study [22]. Key outcome measures of interest for people with dementia (Textbox 2) are cognition and QoL, as previous CST research has shown improvements in these domains [1]. For carers, the key outcome measure (Textbox 3) is QoL as previous iCST research has shown improvement in the QoL of carers [2]. This study additionally includes technology-related scales to assess the usability and acceptability of the iCST app and computer use self-efficacy (Textbox 4). All assessments will take place in the homes of the participants. Wherever possible, 2 researchers will visit the participants in order to interview the person with dementia and his/her carer separately. It will be possible to conduct the assessments over 2 days in case of fatigue or other practical issues such as lack of time. As in the previous iCST trial, the first follow-up at 5 weeks will be included to safeguard data against loss-to-follow-up [2]. The second follow-up will take place at 11 weeks after the baseline as this should be the point that participants in the experimental group will have completed each activity on the iCST app.

Textbox 2. Outcome measures for people with dementia.

<table>
<thead>
<tr>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognition measured using the Alzheimer’s Disease Assessment Scale-Cognitive Subscale [23].</td>
</tr>
<tr>
<td>Quality of life measured using the Quality of Life-Alzheimer’s Disease (QoL-AD) questionnaire [24]. Carers will complete the family version of the QoL-AD.</td>
</tr>
<tr>
<td>Health-related quality of life measured using the EuroQoL 5-dimension questionnaire [25].</td>
</tr>
<tr>
<td>Relationship quality measured using the Quality of the Carer-Patient Relationship questionnaire [26].</td>
</tr>
<tr>
<td>Symptoms of depression measured using the Cornell Scale for Depression in Dementia [27]. Carers will complete this questionnaire as an informant.</td>
</tr>
<tr>
<td>Behavioral disturbances measured using the Neuropsychiatric Inventory and will be rated by the carer only [28].</td>
</tr>
<tr>
<td>Functional abilities measured using the Bristol Activities of Daily Living Scale [29] and rated by the carer only.</td>
</tr>
</tbody>
</table>

Textbox 3. Outcome measures for carers.

<table>
<thead>
<tr>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health-related quality of life measured using the EuroQoL 5-dimension questionnaire [25].</td>
</tr>
<tr>
<td>Anxiety and depression measured using the Hospital Anxiety and Depression Scale [30].</td>
</tr>
<tr>
<td>Relationship quality measured using the Quality of the Carer-Patient Relationship questionnaire [26].</td>
</tr>
</tbody>
</table>

Textbox 4. Technology scales for people with dementia and carers.

<table>
<thead>
<tr>
<th>Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy beliefs in computer/tablet use measured at baseline only using the Computer User Self-Efficacy scale [31].</td>
</tr>
<tr>
<td>Usability and acceptability of the individual cognitive stimulation therapy app measured at the second follow-up and in the experimental group only by using the Questionnaire of Usability and Acceptability [32].</td>
</tr>
</tbody>
</table>

Posttrial Interviews

A small proportion of dyads in the experimental group will be invited to participate in joint semistructured interviews. The purpose of the interviews is to gain additional information on the layout and content of the iCST app, the overall experience of using it as a dyad, and any practicalities surrounding its use in daily life. The interview serves as a complementary data collection method to the weekly telephone support calls and the usability questionnaire, as a semistructured interview generates more in-depth data that otherwise cannot be accessed through quantitative methods only [33]. A discussion guide will be developed, including the key areas mentioned before, to help guide the interview. Each dyad in the experimental group will be invited to participate in the interview upon completion of the study. If they are interested, further details for the person with dementia and carer will be sent. If a dyad agrees to participate, a date for the interview will be set. All interviews will take place in the home of the participants. Written informed consent will be obtained from both participants by the unblinded researcher. The data will be audio-recorded, transcribed, and subsequently stored on a password-protected computer at the University of Nottingham.

End of Study

The second follow-up constitutes the end of the study for participants. At this final visit by the blinded researcher, all participants will be given a £10 (US $13.01) App Store or Google Play store voucher in order for them to download the iCST app once it has been released on the app stores. This will...
be accompanied by an instructional document on how to redeem the voucher and a newsletter containing information on what will happen next, such as making improvements to the app and analyzing and disseminating the results.

**Ethical Considerations**

Ethical approval has been obtained from the Yorkshire and the Humber-Bradford Leeds Research Ethics Committee and NHS Health Research Authority in March 2018 (reference number 17/YH/0405).

**Consent**

People with mild-to-moderate dementia will be recruited in the study and are expected to be able to give informed consent for participation, provided that appropriate care is taken in explaining the research and sufficient time is allowed for them to reach a decision. Written informed consent will be taken at baseline from both the person with dementia and the carer. Since the intervention requires joint participation, it is likely that both participants will consult each other in making their decision. Therefore, it is possible that any individual participant’s decision to either participate or not participate with the research may be influenced by the other participant. However, it is important that individual participants are not forced to make a decision against their will and the researcher will spend as much time as necessary in speaking to the participants individually about the research. It will be made clear to both people with dementia and carers that no disadvantage will accrue in terms of the current care they receive or any future research opportunities if they choose not to participate or withdraw from the study. The consent form will be signed and dated by the participant and the researcher before they enter the study. One copy will be given to the participant and one will be retained at the local study site.

**Adverse Events**

Previous work with CST, maintenance CST, and iCST has not documented any harmful side effects nor any serious adverse events from participating in the intervention activities [1,2,34]. Given that the iCST app is based on the principles of CST and follows a comparable structure to iCST, it is expected that this study will not lead to harmful side effects for either the person with dementia or the carer. Researchers will be made aware of any adverse events during follow-up assessments or the weekly telephone support calls. The trial manager and chief investigator of the study will be informed in case of any adverse events, and they will assess the severity of the adverse event. Serious adverse events include death, illness related to a previous health condition, or hospitalization.

**Data Security and Entry**

Each study site will create their own password-protected spreadsheet containing participant identifiable information and allocation outcome for the dyad. This spreadsheet can only be accessed through a secure NHS Trust computer. After collection of the data from each site, the data will be stored in a secure cabinet at the University of Nottingham. Identifiable information, including the consent forms, will be kept in a separate locked cabinet. After reviewing the data and checking the scoring, it will be entered manually into SPSS version 25 (IBM Corp) for Windows, which will be used for all the analyses.

**Statistical Analyses**

Key feasibility outcomes will be reported through frequencies and will include the number of participants screened, recruited, randomized, and retained through the duration of the trial. Adherence to the intervention will be assessed by calculating the average number of iCST app activities completed by the dyad logged in the weekly telephone calls. The usability and acceptability of the iCST app will be further investigated by examining data from the weekly telephone calls, posttrial interviews, and by calculating scores on the Questionnaire of Usability and Acceptability, with higher scores indicating higher levels of usability and acceptability. Data from the posttrial interviews will be coded and summarized, and they may be analyzed thematically with specialized software if sufficient data have been obtained to reach data saturation [35]. Lastly, outcome measures will be assessed for appropriateness by calculating missing data rates within the measures and across the follow-ups. As this is a feasibility trial and null hypothesis significance testing is inappropriate due to a likely lack of power to detect significant effects of the intervention [36-38], analyses will mainly include descriptive statistics computed for each group and outcome measures, including means, standard deviations, 95% confidence intervals, and effect sizes [12,39]. However, in order to compare the outcomes on each of the questionnaires between the 2 groups, an analysis of covariance will be undertaken, which will help to better understand any trends in the data. All analyses will be based on the intention-to-treat principle in that all available data will be included in the analyses. Rules for missing data will be adapted from the main iCST trial [2]. Data will not be imputed if outcome measures or assessments are missing in full, and imputation (using prorating) will only be used when fewer than 20% of cases are missing on any given measure.

**Results**

Funding for this study was obtained in 2015. Development work for the iCST app lasted 1 year and took place between October 2017 and October 2018 [11]. Data collection for the feasibility RCT started in November 2018; however, due to the COVID-19 pandemic, recruitment had to be terminated prematurely. Data collection was completed in March 2020. Data cleaning, entry, and analyses are underway. The full results of this study are expected in the spring of 2021.

**Discussion**

**Overview of This Study**

The iCST app is the first computerized version of CST, which can be used on touchscreen tablets by people with dementia and their carers. It aims to provide mental stimulation and to facilitate conversation between dyads through the use of interactive touchscreen technology. Based on previous research, it is expected that regular use of the ICST app could potentially lead to improved cognition and QoL for the person with dementia and his/her carer. This is an innovative feasibility
RCT that sets out to evaluate the feasibility of conducting a full-scale RCT with the iCST app compared to a TAU control group and to assist the development of a protocol for a full-scale trial. A range of data will be collected on relevant study-related aspects for a potential full-scale RCT, including the study design and process and the feasibility and usability of the iCST app. Data collection is supported by a mixed-methods approach where quantitative data from questionnaires and analytics will be complemented by qualitative data from telephone calls and interviews with people with dementia and carers. A few challenges in the recruitment process are anticipated, given that the iCST app is only compatible with certain touchscreen tablets and software versions and requires internet access at home, which may lead to the exclusion of dyads not meeting these criteria. These can be overcome by ensuring an appropriate timeframe for recruitment supported by adequate resources in terms of capacity at study sites. The findings from this feasibility RCT will be used to draw recommendations in terms of conducting a full-scale trial and determine which modifications are necessary. This will be done using the Acceptance Checklist for Clinical Effectiveness Pilot Trials, which consist of several trial components ranging from trial design and interventions to randomization and data procedures [40]. It will be used to determine which components of the trial will need amendments and how this can be achieved. In addition to a large-scale RCT with the iCST app, other research activities could consist of an implementation study, which would investigate the cost-effectiveness and the accessibility of the iCST app among people with dementia from varying backgrounds. This could include consultation work with people with dementia and carers in order to explore facilitators and barriers toward accessing technology, the iCST app, and its use.

Conclusions

This study will give insights into the feasibility of conducting a full-scale RCT with the iCST app compared to a TAU control group. The full results of this feasibility RCT, including data on the intervention in terms of usability and adherence and outcome data are expected in early 2021. These results will inform whether a full-scale RCT is feasible and which modifications to the study design and process and intervention are needed.

Acknowledgments

We would like to thank all the research participants who have taken part in this study so far and for providing us with essential feedback on the iCST app. We are also thankful to the research staff at Derbyshire Healthcare NHS Foundation Trust, Lincolnshire Partnership NHS Trust, Northamptonshire Healthcare NHS Foundation Trust, and Nottinghamshire Healthcare NHS Foundation Trust for supporting this study. The research presented in this paper is carried out as part of the Marie Curie Innovative Training Network action, H2020-MSCA-ITN-2015, under grant agreement number 676265.

Conflicts of Interest

MO was involved in the development and publishing of the paper-based CST manuals for dementia. Royalties from the sales of these manuals go to the support of the international CST center at University College London run by Dr Aimee Spector. Royalties from the sales of the iCST app (Thinkability) go to Eumedianet and the University of Nottingham and provide support for ongoing maintenance of the app, including future updates. The remaining authors have no conflicts to declare.

References


Abbreviations

CIRCA: Computer Interactive Reminiscence and Conversation Aid
CST: cognitive stimulation therapy
iCST: individual cognitive stimulation therapy
NHS: National Health Service
QoL: quality of life
RCT: randomized controlled trial
TAU: treatment-as-usual

©Harleen Kaur Rai, Justine Schneider, Martin Orrell. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 08.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Virtual Reality Self-help Treatment for Aviophobia: Protocol for a Randomized Controlled Trial

Jamie Rhiannon Fehribach1,2, BA, MSc; Marieke Bianca Jolien Toffolo1, MSc, PhD; Ilja Cornelisz3,4, MPhil, MSc, PhD; Chris van Klaveren3,4, MSc, PhD; Annemieke van Straten1,2, MSc, PhD; Jean-Louis van Gelder5,6, MA, PhD; Tara Donker1,2,7, MSc, PhD

1Clinical Psychology, Department of Clinical, Neuro- and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
2Amsterdam Public Health Research Institute, Amsterdam, Netherlands
3Methods and Statistics, Department of Education Sciences, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
4Amsterdam Center for Learning Analytics, Amsterdam, Netherlands
5Department of Criminology, Max Planck Institute for the Study of Crime, Security and Law, Freiburg, Germany
6Institute of Education and Child Studies, Faculty of Social and Behavioural Sciences, Leiden University, Leiden, Netherlands
7Laboratory of Biological and Personality Psychology, Department of Psychology, Albert Ludwigs-University of Freiburg, Freiburg, Germany

Corresponding Author:
Tara Donker, MSc, PhD
Clinical Psychology
Department of Clinical, Neuro- and Developmental Psychology
Vrije Universiteit Amsterdam
Amsterdam
Netherlands
Phone: 31 641313555
Email: t.donker@vu.nl

Abstract

Background: Aviophobia (the fear of flying) can greatly impact the daily life functioning of people with the condition. Traditional exposure-based treatment is hampered by the limited availability of airplane practice situations, which is a result of economical and practical concerns. Easily accessible and low-cost virtual reality exposure therapy may address these challenges.

Objective: The purpose of our study is to investigate the effectiveness of ZeroPhobia: Aviophobia (a self-help mobile app–based treatment) in reducing flight anxiety symptoms and depressive and anxiety symptoms. We will also investigate the effects of usage intensity, the sense of immersion, inherent absorption ability, and perceived user-friendliness on the treatment effect.

Methods: Participants (N=114) who are aged 18-64 years and experience at least mild symptoms of aviophobia will be recruited from the general Dutch population and randomized into a treatment group or waitlist control group. By using their own phones and rudimentary mobile virtual reality headsets, participants will receive six modules of psychoeducation and cognitive behavioral therapy, which will include six levels of virtual reality exposure therapy over a period of 6 weeks. Assessments will be conducted at baseline, posttest (ie, after 6 weeks), and 3- and 12-month follow-ups. The primary outcome measure of our study is the Flight Anxiety Situations Questionnaire. The secondary outcome measures include anxiety and depression measures and additional covariates (including usage intensity, the degree of immersion, etc). We will test treatment effectiveness by conducting an intention-to-treat analysis and estimating average treatment effects on the treated. The mechanisms of treatment effect will also be explored.

Results: The study was funded on September 25, 2018. Ethical approval was received on October 11, 2019. Recruitment closed on May 7, 2020.

Conclusions: Our study will further the scientific understanding and clinical implications of technology’s current ability to aid in providing effective, accessible treatment for the fear of flying.

Trial Registration: Netherlands Trial Registry NL70238.029.19; https://www.trialregister.nl/trial/8257.

International Registered Report Identifier (IRRID): DERR1-10.2196/22008

(JMIR Res Protoc 2021;10(4):e22008) doi:10.2196/22008
KEYWORDS

aviophobia; specific phobia; virtual reality; cognitive behavioral therapy; exposure therapy

Introduction

Aviophobia, or the fear of flying, is one of the most common specific phobias. It has a lifetime prevalence rate of approximately 13.2% [1]. An estimated 10%-35% of people from Western Europe and North America experience enough flight anxiety that they avoid flying or do so with great discomfort and fear [2]. This avoidance, which is a common characteristic of specific phobia, not only perpetuates this fear, but also results in interruptions to one’s ability to live life. Being unable to visit family members abroad or having to quit a job that requires air travel are just two examples of how the fear of flying can impact an individual [3]. The high degree of aviophobia’s diffusive impact on day-to-day life is, in part, due to the heterogeneity of the phobia. Aviophobia stems from a range of subfears and reasons for fear [2]. Anxiety toward flying can be rooted in a fear of heights (acrophobia) or a fear of small or inescapable spaces (claustrophobia). Such anxiety can also result from fears of a plane crashing or the inability to have control over a flight experience [2]. Aviophobia, as well as other specific phobias, stands the risk of becoming chronic if left untreated. It is also connected with the comorbid development of other mental health problems [4]. Therefore, effective, accessible, and affordable treatments are necessary.

Specific phobia is most commonly treated via cognitive behavioral therapy with exposure therapy [5]. During exposure therapy, a person is gradually exposed to their feared stimuli, starting from their least to most feared stimulus. Exposure therapy is most often conducted face-to-face with a therapist in real-world (or in vivo) environments [6]. This design however gives rise to several challenges, especially when it comes to aviophobia. For instance, (1) monetary and temporal costs are high due to the need for a therapist and real-world materials for conducting an exposure (e.g., plane tickets, time for traveling to and from the airport or flight destination, etc.) [7,8]; (2) waitlists are longer, again due to the need for a therapist over the course of treatment [9]; and (3) many individuals are too afraid to begin exposure therapy or drop out before the therapy is complete due to fears of exposures [10].

Researchers have already begun to search for new ways to approach these challenges, including the implementation of virtual reality (VR) technology. Unlike traditional methods, VR exposure therapy (VRET) uses computer-generated environments for exposure, thus addressing the costs of travel and stimuli for exposures. Additionally, research has shown that individuals are less averse to beginning VRET and more averse to beginning in vivo treatment due to a fear of confronting feared stimuli in real life [11].

Several studies on the fear of flying that use VRET have demonstrated positive results [12,13]. A meta-analysis found that VRET for the fear of flying was superior to control conditions (g=1.35; P<.001) and classical evidence-based treatments (g=0.35; P=.01) that included interventions that are used to conduct in vivo exposure therapy [14]. Therefore, VRET for the fear of flying is a viable alternative to in vivo treatment. Although VRET addresses many of the challenges in treating patients with aviophobia, most treatments still involve a therapist to guide clients through the program [15]. This means that waitlists for available therapists still impede people who seek treatment. Several VRET studies have demonstrated the potential of technological treatments to be effectively self-guided [16-18]. In these VRET studies, participants followed treatment programs without a guiding therapist. However, they were still required to complete sessions in a lab with an assistant present, making these studies’ results less reflective of real-life experiences.

ZeroPhobia is a VRET program that was designed to investigate whether VRET can in fact be completed at home without a therapist and with basic equipment. This program only requires the user to have a compatible phone (i.e., one that contains a gyroscope) and a mobile VR headset, thus reducing the severity of the aforementioned treatment barriers (temporal and monetary costs, waitlists for an available therapist, and aversion to facing real-life stimuli). Furthermore, it has been reported that ZeroPhobia is effective in treating acrophobia, with an effect size (d) of 1.14 (intention-to-treat [ITT] group compared to a waitlist group) [15].

The primary aim of our study is to investigate whether ZeroPhobia: Aviophobia (a self-help mobile app–based treatment) can effectively reduce flight anxiety symptoms by using the most basic and affordable VR equipment (i.e., a mobile VR headset and participants’ own smartphones) and whether these improvements can be maintained in the long term. This will be confirmed by analyzing changes in flight usage between the pretest and posttest periods and at follow-up time points. We will also test the effects of ZeroPhobia on anxiety and depressive symptoms. Additionally, we aim to investigate whether ZeroPhobia: Aviophobia is user-friendly, to measure app usage intensity, and to determine whether participants felt present and immersed in VR.

Although VR technology’s ability to induce feelings of “being there” (presence) through immersive technology is well known, studies have found conflicting evidence on the importance of presence for successful VRET [8,19-24]. One factor that could influence this relationship is a person’s inherent ability to feel like they have been transported into a fantasy environment. This feeling is known as absorption [24,25]. Studies have suggested that during VRET with low-immersion technologies, an individual’s degree of absorption is related to the efficacy of the treatment [26]. Since ZeroPhobia uses relatively low-immersion technology, absorption may be an important mediating factor for our outcome. Therefore, absorption will be explored.

We hypothesize that participants from the general Dutch population who are randomized into the treatment group will experience a greater reduction in flight anxiety symptom severity.
between the pretest and posttest period compared to those in the waitlist control group. We also hypothesize that this reduction in flight anxiety symptom severity will be maintained at the 3- and 12-month follow-ups. Additionally, we hypothesize that individuals will experience less severe general anxiety and depression after treatment. We further hypothesize that ZeroPhobia: Aviophobia will be rated as user-friendly and that participants will feel immersed in VR environments. Finally, we hypothesize that higher usage intensity, a higher sense of presence, higher inherent absorption ability, and higher perceived user-friendliness will be associated with a reduction in flight anxiety symptom severity at posttest.

Methods

Study Design

Our study is a randomized controlled trial with a superiority design. We will compare the following two conditions: the intervention and waitlist control conditions. The duration of both the treatment intervention and waitlist waiting periods is 6 weeks. After the waiting period, participants who were randomized into the control group will receive 6 weeks of treatment.

A total of 114 participants will be recruited, and each condition group will include 57 participants. Further details are provided in the “Sample Size” section of this paper. Participants in the intervention group will complete web-based questionnaires at baseline, posttest, and the 3- and 12-month follow-up time points. Participants in the waitlist control group will complete web-based questionnaires at baseline, at posttest, and after receiving the intervention treatment. Figure 1 shows a participant flowchart. Multimedia Appendix 1 includes a SPIRIT (Standard Protocol Items: Recommendations for Intervventional Trials) flow diagram.

Our study received ethical approval from the Medical Ethical Committee of Vrije Universiteit Medical Center (registration number: 2019-321). The study is registered in the Netherlands Trial Registry (trial number: NL70238.029.19).
Procedure

Participants will be recruited from the general Dutch population via advertisements on web-based platforms (such as Facebook, Instagram, and webpages that belong to the following fear of flying courses: Stichting Valk and Onbezorgd Vliegen) and offline platforms (ie, national radio, newspaper/magazine articles, and flyers). These advertisements will direct interested individuals to the study website, on which detailed information about the study is available. Those who are interested and believe that they are eligible will be directed to complete a brief web-based response form to express their interest. No incentives were offered to participants other than the opportunity to receive the novel treatment as part of our research.

The research team will then respond to interested individuals by providing additional informational materials and a web-based screening questionnaire to determine eligibility. Those who complete our screening questionnaire but are deemed ineligible will receive an automated email response informing them of...
why they are unable to participate in the study. Potential participants will be reminded via email twice to complete the screener; if potential participants do not provide a response or are no longer interested, they will be removed from our contact list.

Those who are found to be eligible for the study will receive an informed consent form and participant information letter with an enclosed preaddressed return envelope via mail. Participants will receive two reminders (first by email, then by phone) to return the signed consent form. This consent form will ask participants to consider whether they would like to receive information on the results of the overall study, which we will share after the analysis is complete. After returning their signed consent form via mail or email, participants will be provided with our baseline questionnaire. After completing this questionnaire, participants will be included in the study, and an independent researcher will randomly assign participants to either the treatment group or waitlist group by using the Random Allocation Software program (MathWorks). Participants will be informed of which group they are assigned to; those randomized into the waitlist group will be told that they will be contacted again in 6 weeks, and those randomized into the treatment group will be provided with a mobile VR headset and usage instructions via mail. Participants will also be provided with instructions on how to download and install the ZeroPhobia app and a unique unlock code to access the app.

After downloading the app, participants will be able to freely use the ZeroPhobia app at their own pace. The app includes 6 modules that participants can complete. Weekly emails will be sent to remind participants to continue using ZeroPhobia during the treatment period. These emails will be sent out regardless of how much time a participant has spent on using the app. Weekly surveys that measure flight anxiety were not included in this procedure. This allows us to use the most naturalistic approach of applying the intervention and to lighten the burden on an individual. In the final module, participants will be encouraged to engage in real-world exposures (ie, booking an airplane flight or visiting airports). After 6 weeks, participants from both groups will receive an invitation to complete a posttest questionnaire.

Once posttest questionnaires are completed, intervention group participants will be told that they will be contacted again after 3 and 12 months and that they are free to continue using ZeroPhobia if they wish. Control group participants will be invited to begin using ZeroPhobia. During the intervention period, they will also receive weekly reminders. After 6 weeks of ZeroPhobia use, participants will be asked to complete one final questionnaire that only pertains to the primary outcome measure.

**Inclusion and Exclusion Criteria**

To be included in the study, participants must achieve a score of ≥56 on the Dutch version of the Flight Anxiety Situations Questionnaire (FAS), which indicates mild flight anxiety [27,28]. Additional inclusion criteria include an age of between 18 and 64 years, the ability to speak proficient Dutch, the possession of a compatible smart phone (either an iPhone [version 5 or higher] or an Android phone [that operates on Lollipop version 5.1 or higher with screen sizes of 4.7-5.5 inches] with access to the internet), and the provision of informed consent to participate. Exclusion criteria include receiving other psychological treatments for aviophobia at the time of the study or starting, adjusting, or planning to start or adjust psychotropic medication from 3 months prior to the time of recruitment to the planned end time of the study.

**Sample Size**

Our primary outcome measure of aviophobia, the FAS, was used for power calculations. Previous meta-analytical studies on VRET for the fear of flying have found that such therapy has an effect size (Cohen d) of 0.80 [29]. Since ZeroPhobia: Aviophobia is a self-help treatment that involves rudimentary VR glasses instead of standard VRET with a therapist, we used a slightly more conservative effect size (d=0.70) as the starting point for our power calculations. To identify a difference between the experimental and control conditions with a standardized effect size of 0.70 (determined with a two-tailed t test for assessing differences between two independent means), an α value of .05, and a statistical power (1−β) of .80, we require 68 respondents (34 respondents per group). After taking into consideration an expected dropout rate of 40% [15], our overall required sample size is 114 (57 participants per group).

**Randomization, Blinding, and Treatment Allocation**

Randomization will be conducted with the Random Allocation Software program, which uses the block randomization method for conducting a 1:1 allocation and random block sizes of 6, 8, 10, and 12. Randomization will be conducted by an independent researcher. This researcher will generate a randomization list (a list of outcomes) that will be used to assign participants to either the treatment group or waitlist group. The blinding of participants and the research team to who is assigned to each group is not possible due to the nature of our study. However, all measurements will be completed by the participants via a web-based platform, without the presence of the research team.

**Intervention**

ZeroPhobia: Aviophobia is an app-based intervention that consists of six modules that can be completed at the user’s own pace. ZeroPhobia was created with aesthetic in mind, making the graphics and content visually pleasing and logically designed. There is no guide or therapist involved, only a preprogrammed virtual therapist who explains how to use the app and educates the user. Therapeutic content provided by the modules was created by following the manualized protocols of Cognitive Behavioral Therapy for Specific Phobias [30-32]. The app was designed for smartphone use only, because the device must be small enough to double as the screen for VR glasses during VRET.

Each module takes 5-20 minutes to complete, depending on how fast the user wishes to go through a module. Cognitive behavioral therapy content is provided by the virtual therapist and the use simple, 2D animations and voice overs. Module 1 (Achtergrond/Background) describes what flight anxiety is and facts about airplane safety. In this module, participants are also introduced to two faux past ZeroPhobia users, Chloe and Bruno. Both appear across modules in order to share their experiences.
with ZeroPhobia and offer examples of how flight anxiety develops and how it is overcome, thus further educating participants. Module 2 (Je angst te lijf/Facing Your Fears) teaches participants about the anxiety curve and how to create realistic goals to overcome fears of flying.

Module 3 explains how to effectively and safely undergo VRET. After completing this module, the VR levels become unlocked. These levels include detailed 3D animations of the following flight scenarios: checking in at an airport, boarding a plane, finding a seat, taking off, flying, flying with turbulence, and landing. Each of these scenarios involves games, such as finding objects or comforting other nervous passengers, that are led by the virtual therapist. These games help to increase engagement. Participants are able to unlock and progress through VR levels by rating their anxiety with a score of ≤3. Scores are based on a scale of 1-10, with 1 indicating very little anxiety. If a participant completes a VR level and rates their anxiety during the level with a score of >3, a message appears. This message explains that they should practice this level again before moving on to the next level.

Participants will be encouraged to practice with the VRET levels 10 minutes per day and the final three final modules once per week. Before and after finishing a VRET level, participants are prompted to rate their anxiety (“How high was your anxiety at its peak during this level?”). Participants will receive preprogrammed feedback on this rating, which will advise them to either move on to the next VRET level or repeat the level again until their anxiety is lower.

Participants will be encouraged to practice VRET as they complete the final three modules. Module 4 focuses on educating participants about automatic, catastrophic thoughts and identifying their own thoughts. Module 5 expands on this topic by teaching participants about helping thoughts that they can construct to combat catastrophic thinking. Finally, module 6 encourages participants to use their newfound knowledge and begin to practice exposure in the real world. Participants will continue to have access to ZeroPhobia after the test phase is completed and during follow-up. Table 1 provides an overview of all modules, and Figures 2-5 show images of the ZeroPhobia app.

Table 1. Module overview.

<table>
<thead>
<tr>
<th>Module</th>
<th>Learning objective</th>
<th>Additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1: Achtergrond/Background</td>
<td>Understand what the fear of flying is and how it develops</td>
<td>Participants are introduced to two fictional characters who have also followed ZeroPhobia. These characters will provide further examples.</td>
</tr>
<tr>
<td>Module 2: Je angst te lijf/Facing Your Fears</td>
<td>Create goals and learn about the anxiety curve</td>
<td>Participants create goals such as “I would like to be able to take a short flight in a few months with less anxiety than I feel now.”</td>
</tr>
<tr>
<td>Module 3: Exposure</td>
<td>Learn about exposure, the completion of exposure levels, and safety information</td>
<td>After completing this level, the first virtual reality level is unlocked. In order to continue unlocking VR levels, participants must report low levels of anxiety (ie, a score of ≤3 on a scale of 1-10) for the current level.</td>
</tr>
<tr>
<td>Module 4: Rampgedachten/Catastrophic Thoughts</td>
<td>Learn about automatic, catastrophic thoughts are and how to identify such thoughts (ie, their own thoughts)</td>
<td>Participants learn about how automatic, catastrophic thoughts can increase and maintain anxiety. Participants are also encouraged to reflect on how realistic their catastrophic thoughts actually are.</td>
</tr>
<tr>
<td>Module 5: Helpende gedachten/Helping Thoughts</td>
<td>Formulate helping thoughts to combat catastrophic thoughts</td>
<td>Participants must think of reasons for why their catastrophic thoughts are not realistic and the unrealistic thoughts they have about their fears.</td>
</tr>
<tr>
<td>Module 6: De volgende stap/The Next Step</td>
<td>Become inspired to begin practicing in vivo exposures and formulate a fear hierarchy</td>
<td>An example fear hierarchy is created by using one of the fictional characters, who provide recommendations on how to reward oneself for completing different fear hierarchy challenges.</td>
</tr>
</tbody>
</table>
Figure 2. The virtual therapist and guide of ZeroPhobia.

Figure 3. The home screen of ZeroPhobia. For translations of module names, see Table 1.
Primary Outcome Assessment (FAS)
The FAS [27] is a Dutch, 32-item, self-report questionnaire that is frequently used to assess the fear of flying. The response for each item is based on a 5-point Likert scale (1=no fear; 5=overwhelming fear). The questionnaire focuses on measuring how much anxiety is induced by different airport or flight situations. The reliability of the FAS ranges from good to excellent, and the questionnaire is sensitive in distinguishing between people with and without the fear of flying [28]. The total possible score on this measure ranges between 32 and 160. This measure will be used for screening purposes; it has a cutoff score of 56, which indicates mild anxiety [23,24]. The FAS will be used again at posttest and during both follow-up time points.
Secondary Outcome Assessment

**Flight Anxiety Modality Questionnaire**

The Flight Anxiety Modality Questionnaire (FAM) [27] is a Dutch, 18-item, self-report questionnaire. FAM scores are a secondary measurement to FAS scores; they are used to maintain the robustness of results. Similar to the FAS, responses to FAM items are based on a 5-point Likert scale that ranges from 1 (not at all) to 5 (very much so). The reliability of this questionnaire has also been found to be good [28]. Unlike the FAS however, the FAM focuses on the degree to which different anxiety symptoms are experienced in flight situations. This measure will be assessed at baseline, at posttest, and during both follow-up time points.

**Beck Anxiety Inventory**

The Beck Anxiety Inventory [33] is a self-report questionnaire that consists of 21 items on a 4-point Likert scale, which ranges from 0 (not at all) to 3 (severely; “I could barely stand it”). The total possible score ranges from 0 to 63. Scores of 0-9 indicate no anxiety, scores of 10-18 indicate mild to moderate anxiety, scores of 19-29 indicate moderate to severe anxiety, and scores of 30-36 indicate severe anxiety. The questionnaire’s internal consistency is high (0.90-0.94) and its validity is good [34]. For our study, the Dutch translation of the Beck Anxiety Inventory will be used [35]. This measure will be assessed at baseline, at posttest, and during both follow-up time points.

**Patient Health Questionnaire**

The Dutch translation [36] of the 9-item Patient Health Questionnaire [37] will be used to assess depressive symptoms at baseline, at posttest, and during follow-up time points. This questionnaire has good sensitivity (0.71-0.84) and specificity (0.90-0.97) [38].

**Web Screening Questionnaire**

The Dutch Web Screening Questionnaire [39] consists of 15 items and screens for an array of common mental health disorders (eg, depressive disorder, alcohol abuse/dependence, etc). For the purposes of our study, only items that assess panic disorder, agoraphobia, and obsessive-compulsive disorder will be used, as these disorders involve symptoms that overlap with those of flight anxiety. The Web Screening Questionnaire has good sensitivity (0.72-1.00) and specificity (0.44-0.77) [39]. This questionnaire will only be provided at baseline to assess whether the presence of panic disorder, agoraphobia, or obsessive-compulsive disorder influences posttest flight anxiety outcomes.

**Credibility/Expectancy Questionnaire**

The 6-item Credibility/Expectancy Questionnaire (CEQ) [40] will be used to assess self-reported treatment expectations at baseline. The CEQ has high internal consistency and good test-retest reliability [40]. Studies have also found that the Dutch translation of the CEQ is valid [41]. This questionnaire will be used as a complement to the Client Satisfaction Questionnaire (CSQ) to better understand whether the treatment met participants’ expectations.

**CSQ Assessment**

The CSQ [42] complements the CEQ. It measures participants’ satisfaction with ZeroPhobia at posttest. This questionnaire consists of 8 items that assess whether ZeroPhobia fulfilled participants’ expectations and whether participants would recommend ZeroPhobia: Aviophobia treatment to someone else with similar flight anxiety symptoms. Studies have found that the CSQ is both valid and internally reliable (Cronbach α values range between .83 and .93) [39]. The Dutch translation has also been tested and found to be valid [43].

**System Usability Scale**

The System Usability Scale (SUS) [44] consists of 10 items that measure the user-friendliness of the ZeroPhobia app. The response to each item is based on a 5-point Likert scale (1=not in agreement; 5=very much in agreement). Total scores are calculated and converted so that they range from 0 to 100. Bangor et al [44] have provided details on this conversion. Higher scores indicate better usability and higher user-friendliness. SUS items include “I think I would gladly use the ZeroPhobia app regularly” and “I found the ZeroPhobia app unnecessarily complex.” In order for a system to be considered passable, it must receive a score of ≥70. Better products receive scores of ≥80. Only incredibly high-quality products are expected to receive a score of ≥90. Systems that receive a score of <70 are considered inadequate and should be reassessed and improved. The reliability of the SUS is good [44], and the validity of Dutch translation of the SUS is comparable to that of the original measure [45]. The SUS questionnaire will only be provided (at posttest) to the intervention group.

**Igroup Presence Questionnaire**

The 14-item Igroup Presence Questionnaire [19] assesses the realism of VR environments and the feeling of “presence” or immersion in VR environments. Responses to each questionnaire item is based on a 7-point Likert scale that ranges from −3 (fully disagree) to 3 (fully agree). The questionnaire includes items such as “I had the feeling of being surrounded by the virtual world.” This questionnaire has been found to be reliable (Cronbach α=−.73) [19], Schubert et al [19] translated this questionnaire from German to both Dutch and English; the Dutch version will be used for our study. This questionnaire will only be provided (at posttest) to the treatment group.

**Questions on Real-Life Flight Usage**

Single-item questions will be asked at baseline, at posttest, and during both follow-up time points to understand changes in flight usage during the study. Baseline questions will ask participants about whether they have ever flown, how many flights they have taken in the past year, and how long each of these flights lasted. The questions asked during subsequent time points will ask participants to provide similar information. These questions will be adjusted according to the study phase. For instance, at posttest, participants will be asked “how many flights have you taken in the past six weeks?” These questions were formulated for the purposes of our study.
**Ecological Momentary Assessment**

Directly after each VRET session, the app will prompt participants to rate how much anxiety (ie, peak anxiety) they experienced during a level on a scale of 1-10. These ratings will be used to further our understanding of the efficiency of our VR environments and advise a participant to either move on to the next VR environment or to practice again. For example, an anxiety rating of ≥4 indicates that participants are still experiencing moderate to high anxiety during a level.

**Additional Measures**

**Interpersonal Reactivity Index**

The Interpersonal Reactivity Index (IRI) [46] is a self-report questionnaire that was created to measure different types of empathy. The Fantasy subsection of the IRI will be used at baseline to understand the general fantasizing abilities of participants. The IRI-Fantasy subscale consists of 7 items on a 5-point Likert scale (0=this does not describe me; 4=describes me very well). By using this questionnaire, we will be able to investigate a person’s ability to experience emotional absorption or feelings of being transported into a fantasy world. This questionnaire has been found to be both reliable and valid [46]. The Dutch translation of this scale has also been found to be reliable and valid [47].

**Questions on Professional Treatment**

Single-item questions regarding other psychotherapeutic treatments that participants are undergoing during the study (either psychotherapy or prescriptions for psychotropic medication) will be asked during screening, at posttest, and during both follow-up time points. These questions will be asked to ensure that if a participant does start a different treatment, we can account for this properly in our analysis. Table 1 provides an overview of which assessments are conducted for each time point.

**Analyses**

**Covariates and Sample Differences**

Analyses will be conducted with Stata 16 (StataCorp LLC). First, the external validity and generalizability of the experimental sample will be determined by comparing it to the overall sample. Second, we will validate whether randomization was successful by examining whether there are significant differences in background characteristics between the intervention and control groups. To check for selective attrition, we will show how missing data (outcome and input data) are divided across experimental groups. To keep the experimental sample intact, a dummy variable will be constructed to represent the missing covariates of a participant. Missing covariate values will be replaced with their mean values.

**Estimating Treatment Effect**

The treatment effect will be estimated by means of ordinary least squares regression, which will be based on an intention-to-treat (ITT) analysis that uses treatment assignment status as the principal predictor. Missing outcome observations for participants will be imputed by using multiple imputation methods for assessing pretest scores and a set of prespecified background characteristics (gender, age, education, and the severity of symptoms).

Nonrandom missing outcome observations can result in biased point estimates; a treatment effect estimate that is obtained via complete case analysis or the estimation of ITT effects may be a biased estimate of the average treatment effects on the treated (ATT). This issue will be addressed via the Random Forest Lee Bounds procedure [48], which is used to trim the outcome distribution of the group with the highest attrition toward the group with the lowest attrition. By removing observations from the lower (upper) end of the distribution, the lower (upper) bound of the ATT can be estimated via ordinary least squares regression. The potentially biased ATT point-estimate will then lie within the estimated treatment effect bounds.

The ATT will be estimated and corresponding measures of clinically meaningful changes and values in treatment effects on the treated will be explored.

**Mechanisms of Treatment**

Underlying mechanisms of observed differences will be investigated via analyses for understanding factors such as user-friendliness, treatment expectation and satisfaction, usage intensity, and feelings of absorption and presence. The robustness and sensitivity of these analyses will be tested by using different assumptions based on the data.

**Data Monitoring and Management**

All data will be collected by researchers from the Clinical Psychology Department of Vrije Universiteit Amsterdam and kept confidential in accordance with the General Data Protection Regulation. The results of our trial will be released regardless of whether they are statistically significant in terms of answering our research questions. This is in line with the publication statement of the Dutch human research commission (Centrale Commissie Mensgebonden Onderzoek).

**Risk to Participants**

Past research has shown that VRET can be administered without introducing significant safety risks to participants [21,48-51]. Additionally, previous research on the ZeroPhobia app has found that using the app does not result in serious adverse effects (no participants reported that they acquired injuries during VR use or experienced unbearable anxiety levels while using the app) [15].

**Protocol Amendments**

If amendments are made to the protocol as it is written now, these will be reflected in the trial registry after receiving approval from the Medical Ethical Committee. The results of our study will be published in peer-reviewed journals once they are available.

**Results**

Our study received funding on September 25, 2018. Ethical approval was obtained on October 11, 2019. Data collection began on November 6, 2019. Recruitment closed on May 7, 2020, and 146 participants were recruited. On July 2020, participants were in the final stages of treatment or in the 3- or
12-month follow-up phase. Data collection is projected to be completed in July 2021.

**Discussion**

In our study, we aim to understand the effectiveness of our self-help app, Zerophobia: Aviophobia, in reducing flight anxiety symptoms. Traditional treatment for the fear of flying is often not accessible to many patients due to the high costs of both the therapists and the materials (ie, plane tickets) that are necessary for in vivo treatment [7,8]. Moreover, many patients are generally reluctant to face their fear and begin in vivo exposure therapy [11]. Untreated specific phobia is related to a high risk of developing anxiety and depression [4], which increase the overall negative impacts of specific phobia and the need for an easily implemented treatment.

Over the past decade, VRET has become a viable alternative to traditional in vivo treatment, as evidenced by VRET's effectiveness, which is comparable to that of in vivo treatment [14]. The evidence on the potential of automated or unguided self-help VRET for phobias also shows that there is a need to further investigate VRET [16,17]. Our study will provide novel insights into automated VRET and methods for making the treatment of specific phobia more accessible to the general population.

Approximately 60%-80% of individuals with specific phobias do not seek treatment for such anxiety disorders, and approximately 25% of individuals with specific phobias who do seek treatment refuse to undergo treatment due to a fear of the therapeutic procedures [11]. With regard to people with aviophobia, those who are interested in receiving treatment face accessibility issues, which are primarily the result of the high costs in acquiring exposure materials (eg, paying for a plane ticket). Our study is therefore specifically aimed at patients who would normally not receive any treatment. As such, our study compares patients who undergo a novel VRET intervention against patients in a waitlist control group. This design is in line with those of other studies that show that a control group is comparable to intervention groups that undergo treatment as usual for a specific phobia [16]. However, we have included measures for the uptake of all interventions/counterinterventions in the treatment and control groups. In future studies, it would be interesting to compare the effectiveness/cost-effectiveness of regular therapy directly to that of our new VR therapy.

In summary, our study will assess whether an unguided self-help VR app that uses rudimentary VR glasses will be effective in reducing the severity of people’s fear of flying. We will also investigate whether symptoms of depression and anxiety are influenced by Zerophobia: Aviophobia treatment. Additionally, our study will determine whether the user-friendliness of similar apps, the degree of immersion in VR environments, people’s inherent absorption abilities, or usage intensity influence the effectiveness of VRET. Finally, we will assess whether these effects can be maintained in the long term (ie, 3 and 12 months following intervention).

**Acknowledgments**

We extend our gratitude to Bruno de Vos, BA (Studio Barbaar), for designing ZeroPhobia; Doruk Eker, MSc (Orb Amsterdam), for programming ZeroPhobia; and Rufus van Baardwijk, MSc, for providing sounds for ZeroPhobia. Bruno de Vos, Doruk Eker, and Rufus van Baardwijk received financial compensation. We would also like to thank Stichting VALK (VliegAngstbestrijding Leidse universiteit, Koninklijke Luchtvaart Maatschappij Naamloze vennootschap) and Onbezorgd Vliegen for assisting us with recruiting participants to our study.

**Authors’ Contributions**

JRF drafted the manuscript. CVK and IC wrote the Covariates and Sample Differences, Estimating Treatment Effect, and Mechanisms of Treatment sections. All coauthors contributed to all sections of the manuscript and conducted a critical review. TD obtained funding and designed the study. TD and JLVG created the VR app. JRF recruited and screened participants. All authors read and approved the final version of the manuscript.

**Conflicts of Interest**

TD and JLVG developed the ZeroPhobia app (the app that is analyzed in our study) in collaboration with Vrije Universiteit. ZeroPhobia is intended for commercial release. Hence, TD and JLVG will not be involved in data analysis or any decisions that are related to the publication of our findings. The other authors declare that they have no competing interests.

Multimedia Appendix 1

SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) flow diagram. Only participants who were randomized into the treatment group were assessed with these measures.

[PDF File (Adobe PDF File), 61 KB - resprot_v10i4e22008_app1.pdf ]

**References**


Abbreviations

ATT: average treatment effects on the treated
CEQ: Credibility/Expectancy Questionnaire
CSQ: Client Satisfaction Questionnaire
FAM: Flight Anxiety Modality Questionnaire
FAS: Flight Anxiety Situations Questionnaire
IRI: Interpersonal Reactivity Index
ITT: intention-to-treat
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SUS: System Usability Scale
VALK: VliegAngstbestrijding Leidse universiteit, Koninklijke Luchtvaart Maatschappij Naamloze vennootschap
VR: virtual reality
VRET: virtual reality exposure therapy

Please cite as:
Virtual Reality Self-help Treatment for Aviophobia: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2021;10(4):e22008
URL: https://www.researchprotocols.org/2021/4/e22008
doi:10.2196/22008
PMID:33843605

©Jamie Rhiannon Fehribach, Marieke Bianca Jolien Toffolo, Ilja Cornelisz, Chris van Klaveren, Annemieke van Straten, Jean-Louis van Gelder, Tara Donker. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 12.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
De-Implementing Opioid Use and Implementing Optimal Pain Management Following Dental Extractions (DIODE): Protocol for a Cluster Randomized Trial

D Brad Rindal1*, DDS; Stephen E Asche1*, BA, MA; Jan Gryczynski2*, PhD; Sheryl M Kane1*, BS; Anjali R Truitt1*, PhD, MPH; Tracy L Shea1*, BS; Jeanette Y Ziegenfuss1*, PhD; Robert P Schwartz2*, MD; Donald C Worley1*, DDS; Shannon G Mitchell2*, PhD

1HealthPartners Institute, Bloomington, MN, United States
2Friends Research Institute Inc, Baltimore, MD, United States
*all authors contributed equally

Corresponding Author:
D Brad Rindal, DDS
HealthPartners Institute
8170 33rd Ave So
MS21112R
Bloomington, MN
United States
Phone: 1 952 967 5026
Email: donald.b.rindal@healthpartners.com

Abstract

Background: Overdose deaths from prescription opioid analgesics are a continuing crisis in the United States. Opioid analgesics are among the most frequently prescribed drugs by dentists. An estimated 5 million people undergo third-molar extractions in the United States each year, resulting in postoperative pain. Studies show that, in most cases, the combination of ibuprofen and acetaminophen is an effective alternative to commonly prescribed opioid analgesics for the management of postextraction pain. Nevertheless, many dentists routinely prescribe opioids after dental extractions.

Objective: We describe the rationale, design, and methods for a randomized trial of interventions designed to de-implement opioid prescribing by dentists while implementing effective nonopioid analgesics following dental extractions.

Methods: Using a prospective, 3-arm, cluster randomized trial design with dentists as the unit randomized and patient-level prescribing data as the primary outcome, we will compare different strategies to reduce the reliance on opioids and increase the use of alternative pain management approaches utilizing information support tools aimed at both providers and their patients. The study will test the efficacy of 2 interventions to decrease opioid prescribing following dental extractions: clinical decision support with (CDS-E) and without patient education (CDS). Providers will be randomized to CDS, CDS-E, or standard practice. Patient-level outcomes will be determined via review of comprehensive electronic health records. We will compare study arms on differential change in prescribing patterns from pre- to postimplementation of the intervention. The primary outcome of interest is a binary indicator of whether or not the patient received an opioid prescription on the day of the extraction encounter. We will also examine recommendations or prescriptions for nonopioid analgesics, patients’ perceptions of shared decision making, and patients’ pain experiences following the extraction.

Results: The HealthPartners Institutional Review Board approved the study. All study materials including the CDS and patient education materials have been developed and pilot tested, and the protocol has been approved by the National Institute of Dental and Craniofacial Research. The intervention was implemented in February 2020, with 51 dentists who were randomized to 1 of the 3 arms.

Conclusions: If the intervention strategies are shown to be effective, they could be implemented more broadly in dental settings with high levels of opioid prescribing.

Trial Registration: ClinicalTrials.gov NCT03584789, https://clinicaltrials.gov/ct2/show/NCT03584789

International Registered Report Identifier (IRRID): DERR1-10.2196/24342
The United States is experiencing an epidemic of prescription drug overdose deaths, with deaths associated with prescription pain relievers of particular concern [1]. Drug overdose has become a leading cause of accidental death in the United States [2]. Between 2000 and 2015, the rate of deaths from drug overdoses increased 137%, including a 200% increase in the rate of overdose deaths involving opioids (opioid pain relievers and heroin) [3]. Unnecessary opioid prescribing is one of the factors driving this epidemic.

The Centers for Disease Control and Prevention strongly recommends fundamental changes in prescribing practices in order to address this public health emergency [4]. Opioid analgesics are among the most frequently prescribed drugs by dentists [5]. For example, an estimated 5 million people undergo third-molar extractions in the United States each year, resulting in postoperative pain, swelling, and discomfort, even when surgical complications are not present [6]. As part of postoperative dental pain management following third-molar (ie, wisdom teeth) surgery, dentists often introduce patients under 25 years of age to prescription opioids for the first time, which can lead to longer-term substance use [7]. In addition, when opioids go unused by the patient, they are sometimes taken by family members or friends, which can also lead to misuse or abuse [8].

Comprehensive reviews of research evidence have concluded that nonsteroidal anti-inflammatory drugs (NSAIDs) are remarkably effective analgesics for relieving postoperative dental pain and that opioid analgesics have a high incidence of adverse effects [9-11]. Evidence concludes that the combination of the NSAID ibuprofen and acetaminophen (APAP) provides analgesia that is at least equivalent to that of commonly prescribed opioid combination formulations [12]. Thus, NSAID+APAP provides a viable and evidence-based pain management alternative to prescription opioids [12]. Nevertheless, most dentists report that they prescribe opioid medications such as hydrocodone or oxycodone following third-molar extractions [13,14]. Many dentists appear to underestimate the immediate risks and the long-term harms associated with prescription opioids, even as they overestimate opioids' therapeutic benefits [15]. Strategies to support dentists in de-implementing their overreliance on prescription opioids following dental extractions in favor of safer alternatives are urgently needed.

Various strategies are being examined to reduce reliance on and misuse of prescribed opioids for management of pain following medical procedures [16,17]. Self-management skills such as patient education, decision making, and forming a patient-provider partnership have long been associated with chronic disease management [18] but can be effective components of acute pain management as well. While patient-provider shared decision making is considered the preferential choice when no clear treatment option is optimal [19], involving the patient in the decision process has distinct advantages with respect to medication adherence as well [20]. Collaborative decision making involving the patient and the dentist holds the potential to decrease reliance on opioid medications in favor of recommended alternatives [21]. The study team will utilize this strategy to engage patients in the decision-making process related to the choice of analgesics to manage their pain following dental extractions.

A well-designed clinical decision support (CDS) system can support dentists in providing optimal pain management for patients without resorting to opioids when a safer alternative would suffice. CDS provides pertinent clinical information to the dentist. It can provide evidence-based information and guidance in the form of prompts and reminders to inform clinical decisions about prescriptions. CDS also provides the advantage of ensuring fidelity of the implementation strategy.

The CDS incorporates novel design features that will save significant time, ensure good fit into the dental care workflow, and facilitate the delivery of personalized care by providing the patient’s relevant medical history, informing treatment decisions, and promoting more evidence-driven pain management following tooth extractions. This implementation project has the potential to drive a major improvement in prescribing practices and the management of pain following tooth extractions in the field of dentistry. As new strategies for pain management are ready to be introduced into clinical practice, they can be readily integrated into this platform in a timely and transparent manner [22].

The primary objective of this project is to test the ability of CDS to support the de-implementation of opioid analgesics and to advance the use of nonopioid (NSAID+APAP) analgesics to manage pain following dental extractions. Using a prospective, 3-arm, cluster randomized trial design with dentists as the unit randomized and patient-level prescribing data as the primary outcome, the study team will compare different strategies to reduce the reliance on opioids and increase the use of alternative pain management approaches utilizing information support tools aimed at both providers and their patients (Figure 1). The primary objective is to test the efficacy of 2 interventions (CDS with and without patient education), compared to treatment as usual to decrease opioid prescribing for dental extractions. Drawing from encounter-level data from the electronic health record (EHR), we will compare study arms on the percentage of extraction encounters with an opioid prescribed.
Secondary objectives are to (1) test the efficacy of 2 interventions (CDS with and without patient education) compared to standard practice to increase exclusive nonopioid pain management for dental extractions, (2) compare the degree to which each of the 3 study arms (CDS with and without patient education and standard practice) facilitates shared provider and patient decision making concerning pain management options for dental extractions, and (3) explore whether the study interventions lead to differences in patient experiences with postextraction pain.

Methods

Study Design

This study is a prospective, 3-arm, cluster randomized trial (Phase III) in which up to 60 dentists (accrual goal: n=51) practicing at HealthPartners dental clinics will be randomized in a 1:1:1 allocation ratio to either standard practice (SP) or 1 of 2 CDS arms (CDS or CDS-E). Patients (up to n=8400; accrual goal: n=6900) will be exposed to the study arm to which their dentist was assigned. Opioid prescribing measured at the encounter level will serve as the primary outcome. Additional outcomes will also be measured at the patient level, including a recommendation or prescription for nonopioid analgesics as
indicated in the EHR and patient-reported outcomes concerning shared decision making and pain levels as identified through patient surveys.

It was impractical to randomize at the patient level for a study that equips providers with CDS. Instead, this cluster-randomized trial design randomizes dentists to 1 of 3 study arms. This design maximizes the number of randomized units (ie, providers) as compared to clinic-level randomization. This design also solves the practical issue of providers who deliver care across several clinics and different study arms that would occur in clinic-level randomization. The institutional review board at HealthPartners Institute reviewed and approved the protocol for this study.

Study Site
HealthPartners is the largest consumer-governed, nonprofit health care organization in the country, providing care, coverage, research, and education to improve health and well-being in partnership with its members, patients, and community. The organization includes a multispecialty medical and dental group practice that includes 75 dentists and 24 dental clinics in Minnesota. For the past 18 years, the medical and dental clinics have fully implemented EHRs with integrated diagnosis codes. The HealthPartners Institute utilizes these systems to conduct research.

Eligibility
Study-eligible providers are practicing dentists including oral surgeons at HealthPartners with a current Drug Enforcement Agency license who performed permanent tooth extractions at a minimum of 6 encounters in the baseline period 1 year prior to implementing the intervention. Study-eligible dental patients are aged 16 years and older and had a permanent tooth extraction performed by a study-eligible HealthPartners provider. Patients are excluded if they have opted out of research at HealthPartners.

Recruitment
Implementation of the CDS is considered a HealthPartners Dental Group initiative. The project is being implemented with the aim of reducing opioid prescribing. Consent has been provided by dental leadership in accordance with HealthPartners practices. Providers randomized to the CDS or CDS-E group will be exposed to the CDS, which is integrated into the EHR. Patients who receive care from providers randomized to the CDS-E arm will also be exposed to study-specific patient education and receive additional educational resources. Thus, as part of the study design, a new section of the EHR is activated for providers in the CDS and CDS-E arms, whereas these EHR resources are left dormant for providers in the SP arm. For the primary study objective, the study does not consent providers or patients.

For the survey-based secondary study objectives, a subset of patients undergoing tooth extractions will be recruited for a telephone survey after the index dental visit. Trained survey administrators will ask these patients to consent verbally for participation via telephone. Stratified sampling will be used to ensure roughly equal representation of patient sample sizes in each study arm and to ensure that patients linked to each provider are sampled for surveys. Sampling weights will be used in the analysis to re-weight the sample results to resemble the population of patients receiving extractions.

Measures to Minimize Bias
The randomization of providers to the study arms will utilize stratification on provider type (oral surgeon vs other) and level of provider opioid prescribing during the baseline period. Providers will be randomly allocated 1:1:1 through a computer-generated program to SP, CDS, or CDS-E. The randomization will be conducted by the study statistician and use a provider identification number with no recognizable meaning, ensuring that the study statistician is blind to provider identity (name). The crosswalk containing the provider identification number, provider name, and study arm will be kept in a secure file and accessible only by the study programmer. In order to minimize contamination across study arms, only providers in the CDS or CDS-E arms will be able to access the CDS. Similarly, patients who receive care from providers in the CDS-E arm will have access to supplemental patient education. As such, both groups of providers will know their assignment. The study team will not disclose the study’s purpose, objectives, or outcomes measures directly to patients until after the encounter or after the participant has been surveyed about their experience, if selected.

Patients will be exposed to the intervention to which their provider has been randomized. At each patient encounter, the CDS web application will collect patient-level data from the EHR (eg, medical conditions, current medications, social history). Providers assigned to the CDS or CDS-E arms will be able to access the information using the CDS tool. Providers assigned to the SP control arm will continue to use their usual methods to assess and address pain management for dental procedures, without any effort from the study to influence their prescribing of analgesics following the dental extraction. They will not be shielded from other outside influences.

At the index encounter (the initial patient visit for a dental extraction where all the study criteria are met), the patient is assigned a unique study identifier that is used to link patient encounter data over time. All index and subsequent encounter data for eligible patients are stored in a limited de-identified analysis dataset. The index encounter is the only patient encounter utilized in the primary analysis. Staff administering the patient survey will not know which arm providers, and thereby survey respondents, have been assigned.

Study Intervention
The study intervention follows the theoretical framework outlined in Figure 2.
Figure 2. Theoretical framework.

**SP**
Providers will not receive point-of-care CDS, serving as the control group representing usual care. Patients will receive usual dental care.

**CDS**
Providers will receive point-of-care CDS to guide pain management recommendations and prescribing for patients who receive dental extractions (see Multimedia Appendix 1). The CDS will highlight (1) potential medication interactions between the patient’s current medications and commonly recommended analgesics for patients following dental extractions, (2) relevant health conditions and identified substance use disorders that may impact pain management strategies, and (3) automated access to the state’s Prescription Monitoring Program. The decision support identifies these potential medication interactions and medical conditions for all analgesic classes (APAP, NSAID, opioid). Patients will receive usual dental care.

**CDS-E**
Providers will receive point-of-care CDS to guide pain management recommendations and prescribing related to dental extractions. Providers will receive the same point-of-care CDS as in the CDS arm (see Multimedia Appendix 1). In addition, patients will receive an educational handout by front desk staff when checking in for the extraction procedure (see Multimedia Appendix 2 and Multimedia Appendix 3). The handout compares opioid and nonopioid pain medications with respect to their risks, benefits, and effectiveness for managing pain. This handout intends to initiate a conversation between the patient and provider about patient needs, goals, concerns, and pain management preferences. This handout also includes information about other pain management after a dental extraction. This handout intends to normalize the experience of some discomfort, help to set realistic expectations, and help patients decide whether to contact their provider for additional help to manage their pain after extraction. Patients will receive usual dental care.

**Clinical Decision Support Design**
The CDS is designed to bring all available clinical information into one interface, saving time searching for the relevant information that would inform the best analgesic options for that individual patient. The CDS identifies potential drug interactions and side effects related to the medications the patient is taking and the potential analgesic options. The CDS messages provide this information to the dentist so it is considered in the analgesic prescribing decision. The messages were informed by information from Micromedex [23], which is considered a highly reliable and current evidence-based source. Medical conditions also need to be considered. The relevant conditions that should be considered when determining the analgesic to prescribe were identified and reviewed by medical experts in the related specialties to ensure that accurate information was provided. The record is also examined for substance use, a history of a substance use disorder, a history of overdose, and a naloxone prescription. Access to the prescription drug monitoring programs is also contained within the CDS. The messaging in the CDS is brief, with the intent of informing a discussion with the patient about why this analgesic recommendation is personalized to them while considering both safety and effectiveness.

**Patient Education Development**
A 9-question survey was completed by 232 patients or plan members through the HealthPartners myVoice panel, providing insights regarding pain expectations and beliefs and preferences for pain management following dental extractions. Findings were used to develop pain management–focused messaging for the enhanced intervention arms of the trial.

**Clinic Workflow Considerations**
In order to optimally situate the timing of the CDS in the EHR, study team members conducted 15 observations of dental extractions with providers delivering standard care using a structured checklist during the study preparation phase, with a focus on provider-patient interactions as well as provider use of the EHR as it relates to analgesic prescribing for extractions.

**Fidelity Monitoring**
Fidelity is monitored by tracking whether the CDS was opened during the clinical encounter. The CDS link is highlighted in the EHR when a treatment plan exists, reminding the provider to open the CDS. However, a treatment plan may not exist for every extraction, for example if the treatment was not determined prior to the visit. We are able to determine if the CDS was opened by the provider, as well as whether a treatment
plan existed. Monitoring these data points informs whether further provider training is needed to improve fidelity.

We also monitor whether the patient education material was printed for all eligible patients. These data inform whether further training of support staff is needed to ensure fidelity.

Data Sources and Endpoints

A data pull from the EHR at the end of the study will serve as the data source for the primary and secondary objectives (see Table 1). Specifically, prescriptions for opioids medications and nonopioid analgesic medications will be available from the EHR. This includes recommendations by the provider to use over-the-counter nonopioid analgesic medications. The EHR will also provide information on prescription quantities, patient demographics, and the dentist conducting the extraction.

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Endpoints</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Differential pre- to postintervention change by study arm in the percentage of extraction encounters with an opioid prescribed on the day of the extraction encounter</td>
<td>EHR²</td>
</tr>
<tr>
<td>Secondary</td>
<td>Differential pre- to postintervention change by study arm in the percentage of extraction encounters at which a provider prescribed or recommended nonopioid analgesics (ibuprofen, naproxen, aspirin, or acetaminophen) and did not prescribe opioids on the day of the extraction encounter</td>
<td>EHR</td>
</tr>
<tr>
<td></td>
<td>Study arm comparison of the mean of the patient-reported shared decision-making composite score (composite of 3 components concerning management of postextraction pain options: effort to explain, listen, and personalize), 3-6 days after the extraction encounter</td>
<td>Patient survey 3-6 days following the extraction encounter</td>
</tr>
<tr>
<td></td>
<td>Study arm comparison of the average patient-reported pain following the extraction, 3-6 days after the extraction encounter</td>
<td>Patient survey 3-6 days following the extraction encounter</td>
</tr>
</tbody>
</table>

²EHR: electronic health record.

A survey conducted within 3-6 days of the extraction encounter in a sample of patients with extractions will be the data source for the remaining 2 secondary objectives concerning patient reports of shared decision making and pain. The survey contains 17 items and is designed to be completed over the phone in less than 5 minutes. Survey constructs include overall satisfaction with the visit, shared decision making, confidence in ability to manage postextraction pain, pain experienced in the days following extraction, and strategies used to manage pain. Questions are based on items with known psychometric properties where available, such as CollaboRATE for shared decision making [24] and the numerical pain rating scale. Demographic questions are based on standard items from Federal Surveillance Surveys such as the Behavioral Risk Factor Surveillance Survey. Where concepts were sought without existing available questions or scales, survey questions were written using best practices for survey questions in order to minimize measurement error and respondent burden [25].

Patient survey data will be collected via phone survey by the research organization’s Center for Evaluation and Survey Research and managed using research electronic data capture (REDCap). REDCap is a secure, web-based application designed to support data capture for research studies, providing (1) an intuitive interface for validated data entry, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for importing data from external sources [26].

Safety Assessments

The study does not include objectives or endpoints concerning safety. However, the monitoring activities will be conducted to assess possible harm.

Sample Size and Power

EHR data from 2018 indicate there were 6900 unique patients aged 16 years and older with a permanent tooth extracted by 51 dentists in the baseline period and also an estimated 6900 unique patients aged 16 years and older with a permanent tooth extracted by 51 dentists during the intervention period. The first visit for a patient in a specific time period (baseline, intervention) at which a patient is eligible for the study serves as the index encounter for the patient to be utilized for the analysis. The study analysis is expected to have at least this many patients. Fully 40% of encounters in the baseline period with patients having a permanent tooth extraction included a prescription for opioid use, and the provider-level opioid use intraclass correlation was 0.3. Assuming the use of a generalized linear mixed model, with an alpha of .05 and 2-sided tests, the planned analysis for the primary objective can detect a differential raw reduction of 23% from pre- to postintervention opioid prescribing when comparing CDS or CDS-E patients (40% with opioid prescriptions pre-interventional to 15%
postimplementation) to SP patients (40% with opioid prescriptions pre-implementation to 38% postimplementation) with 80% power.

Statistical Analysis Plan

Baseline characteristics of study participants will be summarized with mean and standard deviation for interval data and proportions for categorical data. General and generalized linear mixed models will be used to test study arm differences in endpoints. Models will include terms for study arm, time (for models assessing change from pre- to postintervention), and their interaction. Covariates to be included in models will be described in the Manual of Procedures. A random intercept for provider will be included in models to account for the cluster-randomized design with providers as the unit randomized. Differential change will be compared for each study arm via a series of planned contrasts. Study arm contrasts will be tested at an alpha of .05, and all tests will be 2-sided. Model-predicted means and proportions along with 95% confidence intervals will be used to assess the magnitude, direction, and precision of intervention effects.

Results

The study was implemented in February 2020. Initial monitoring indicates that activation of the CDS is very high when an alert is present. The alert is activated by a treatment plan for an extraction. We know that a treatment plan may not exist for every visit because the provider and patient may not make the decision to extract the tooth until the day the tooth is actually removed. As providers become more familiar with the CDS and start to make small changes in their workflow within the EHR, the potential exists for higher levels of utilization.

One month after implementing the study in the clinics, enrollment was halted due to the COVID-19 pandemic. As public health strategies were implemented to reduce the transmission of the virus, federal, state, and professional organizations recommended that dental clinics postpone any nonemergent or elective dental care. In mid-March, HealthPartners implemented these recommendations, and they continue to be in place at the time of this submission. Once dental practice is able to resume, questions remain regarding what changes in operations and equipment will be needed to prevent the spread of COVID-19. We anticipate these changes will have some impact on patient flow and subsequent study accrual rates.

We continue to see changes in opioid prescribing patterns over time. HealthPartners has taken actions to reduce prescribing. These include reducing the default number of opioid tablets that can be prescribed and the creation of opioid guidelines. The Minnesota Legislature and insurance plans have taken additional actions that provide additional regulatory oversight and monitoring of provider prescribing of opioids. These actions are all good if they lead to appropriate reductions in opioid prescribing but may limit our ability to test our intervention strategies, as these strategies are being used within a broader institutional, state, and national context of increased attention on reducing opioid prescriptions.

Discussion

Overreliance on opioids across the health care system contributes to the ongoing opioid use disorder and overdose crisis in the United States. Opioids continue to be used in dentistry following tooth extractions. While opioids are clinically appropriate in some circumstances, evidence shows that, in many cases, nonopioid analgesics are as effective and safer for dental extractions. Strategies are needed that can help the field of dentistry to de-implement the routine use of opioids for dental extractions. CDS tools have demonstrated efficacy at improving various health outcomes in medicine and are showing similar potential in dentistry [27]. A CDS that saves time by bringing all relevant patient information and current evidence to the provider in one interface has the potential to improve clinical decision making as well as patient outcomes. This trial examines both the potential for CDS to improve dentists’ prescribing and the potential additive effect of providing patients with relevant analgesic safety and efficacy information prior to the dental appointment when prescribing decisions are made. The study could highlight effective strategies for de-implementing opioids, improving patient care, and protecting public health. We plan to disseminate the results in relevant journals and through channels made available through the HEAL Initiative.

Acknowledgments

Research reported in this publication was supported by the National Institute of Dental & Craniofacial Research (NIDCR) of the National Institutes of Health (NIH) under Award Number U01DE027441. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

Author RPS has consulted for Verily Life Sciences. Unrelated to the present study, author JG has received research funding from Indivior (paid to his institution and including project-related salary support) and is part of owner of COG Analytics. The other authors have no conflicts of interests to report.

Multimedia Appendix 1
Clinical decision support. [PNG File , 78 KB - resprot_v10i4e24342_app1.png ]

https://www.researchprotocols.org/2021/4/e24342
Multimedia Appendix 2
Creating a plan to manage pain.

[PDF File (Adobe PDF File), 53 KB - resprot_v10i4e24342_app2.pdf]

Multimedia Appendix 3
Caring for yourself after having a tooth removed.

[PDF File (Adobe PDF File), 70 KB - resprot_v10i4e24342_app3.pdf]

Multimedia Appendix 4
CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 14446 KB - resprot_v10i4e24342_app4.pdf]

References

Abbreviations

- APAP: acetyl-para-aminopheno (aka paracetamol or acetaminophen)
- CDS: clinical decision support
- CDS-E: clinical decision support with patient education
- EHR: electronic health record
- NSAID: nonsteroidal anti-inflammatory drug
- REDCap: Research Electronic Data Capture
- SP: standard practice

©D Brad Rindal, Stephen E Asche, Jan Gryczynski, Sheryl M Kane, Anjali R Truitt, Tracy L Shea, Jeanette Y Ziegenfuss, Robert P Schwartz, Donald C Worley, Shannon G Mitchell. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 12.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

Effects of a Digital Mental Health Program on Perceived Stress in Adolescents Aged 13-17 Years: Protocol for a Randomized Controlled Trial

Eliane M Boucher1, PhD; Haley E Ward1, BA; Julia L Stafford1, BSc; Acacia C Parks1, PhD
Happify Health, New York, NY, United States

Corresponding Author:
Eliane M Boucher, PhD
Happify Health
51 East 12th Street, 5th Floor
New York, NY, 10003
United States
Phone: 1 432 258 5233
Email: eliane@happify.com

Abstract

Background: Stress is an important transdiagnostic risk factor in adolescence and predicts a host of physical and psychological problems in adolescence and adulthood. Adolescence is also a developmental stage in which people may be more sensitive or reactive to stress. Indeed, research has shown that adolescents report high levels of stress, particularly when enrolled in school. However, adolescents report engaging in few, if any, stress management techniques. Consequently, the development of effective programs to help address adolescent stress is particularly important. To date, most stress management programs for adolescents are delivered within schools, and the evidence for such programs is mixed. Furthermore, most of these programs rely on traditional stress management techniques rather than incorporating methods to address the underlying negative cognitive processes, such as rumination, that may contribute to or exacerbate the effects of perceived stress.

Objective: The aim of this study is to test the short-term effects of a digital mental health program designed for adolescents aged 13-17 years on perceived stress and rumination.

Methods: This is a randomized controlled trial in which adolescents between the ages of 13 and 17 years, with elevated levels of perceived stress and brooding, will be randomly assigned to complete 8 weeks of a digital mental health program (Happify for Teens) or to a corresponding wait-list control group. The study will take place over 3 months, including the 8-week intervention period and 1-month postintervention follow-up. The primary outcome, perceived stress, along with secondary and exploratory outcomes (ie, brooding, optimism, sleep disturbance, and loneliness) will be assessed via self-report at baseline, 4 weeks, 8 weeks, and 12 weeks to compare changes in these outcomes across conditions.

Results: Recruitment is expected to begin in the second quarter of 2021, with a target sample size of 800 participants (400 per condition). Participants will begin the study as they are recruited and will finish in waves, with the first wave of data expected 8 weeks after recruitment begins and the final wave of data expected by the end of the third quarter of 2021.

Conclusions: Although school-based stress management programs for adolescents are common, research suggests that they may be limited in their reach and more effective for school-based stress than other types of stress. This trial will be one of the first attempts to examine the potential benefits of a digital mental health program on adolescents to address stress along with negative cognitive processes such as rumination. If successful, this would help introduce a more scalable alternative to school-based programs that offers adolescents greater privacy while also providing insight into novel ways to target adolescent mental health more generally.

Trial Registration: ClinicalTrials.gov NCT04567888; https://clinicaltrials.gov/ct2/show/NCT04567888
International Registered Report Identifier (IRRID): PRR1-10.2196/25545

(JMIR Res Protoc 2021;10(4):e25545) doi:10.2196/25545

KEYWORDS
digital intervention; adolescents; stress management; mental health; mobile phone
Introduction

Background

In the United States, approximately 16.5%, or 7.7 million, adolescents have at least one mental health disorder [1]. Worldwide, approximately 10%-20% of adolescents struggle with mental health problems [2]. Importantly, the rates of mental illness have increased among adolescents in recent years [3] and at steeper rates than among adults [4]. Given that the onset of many mental health disorders occurs in childhood and adolescence, which often continue into adulthood [5,6], addressing mental health in adolescence is a key factor for reducing the prevalence of mental health disorders in adulthood. Despite several public health efforts to increase access to mental health services for youth, service utilization by adolescents with mental health disorders remains low [7]. Only 50% of American adolescents with mental illness seek mental health treatment [1], and only 36.2% actually receive treatment [7]. The number of adolescents receiving treatment for internalizing disorders, such as anxiety and depression, is even lower [7,8]. Paradoxically, national trends indicate that, overall, adolescent use of outpatient mental health services, including both psychotherapy and psychotropic medications, has increased over time [9]. However, this increase is primarily attributed to adolescents with less severe or no mental health impairment [9]. This is problematic given the shortage of mental health service providers for children and adolescents, particularly in rural and low-income areas [10]. The increased use of these services by adolescents in general likely places a strain on the available services for youth, and, in turn, adolescents with more serious mental health disorders may have more trouble receiving necessary treatment.

Mental Health Prevention Programs for Adolescents

Although discussions of mental health often focus on dysfunction, more recently, there has been a call for a definition of mental health that considers more than the mere absence of mental illness [11]. For example, Keyes [12] argued for a distinction between languishing, or the absence of mental health, and flourishing, or the presence of mental health. In other words, mental health also encapsulates positive or protective features, including well-being, optimal prosocial functioning, emotion regulation, empathy, and resilience [11,13]. Consequently, addressing mental health concerns in adolescents should also include efforts to promote flourishing, as this may help to reduce the likelihood of future dysfunction [13].

However, there has been more attention on mental health treatment in youth than on prevention [2]. Given the strain on mental health services for youth from adolescents with less severe or no impairment, interventions focused more on promoting flourishing and preventing future dysfunction and impairment are important [13]. Such programs can be universal, which are directed at all adolescents and often offered in schools, or targeted, in that they focus on adolescents who are at risk of more severe impairment [2]. Both universal and targeted programs can help address the growing interest in mental health services by adolescents while reserving more intense treatment programs for youth with mental illness, for whom prevention programs are not sufficient [13]. In particular, interventions that target transdiagnostic risk factors, which predict the onset and maintenance of multiple disorders [14], are likely to have the greatest effect in terms of preventing future dysfunction [15,16].

Research suggests that these prevention programs have both concurrent and long-term benefits for mental health, particularly when they target specific risk and protective factors for adolescent mental health [2,17]. For example, universal resilience-based interventions delivered in schools reduce internalizing symptoms in adolescents aged 11-18 years [18]. In another study, students who were offered a universal, positive psychology intervention reported significant reductions in distress, anxiety, and depressive symptoms, whereas students who were not offered this intervention reported increases in all these outcomes [19]. Targeted prevention programs based on cognitive behavioral therapy (CBT) have also been shown to have short-term effects on depressive symptoms [20,21] and improve self-efficacy and academic achievement in adolescents [22].

Stress as a Risk Factor in Adolescence

Over the past decade, researchers have become increasingly interested in adolescent stress as a transdiagnostic risk factor because of its prevalence and its association with internalizing and externalizing disorders [23-25] as well as a host of other negative consequences, including poorer physical and academic outcomes. Adolescence is also a developmental stage where individuals may be particularly sensitive to stress due to shifts in hypothalamic-pituitary-adrenal axis reactivity, leading to more intense hormonal responses to stressors, particularly in older adolescents [26]. Although animal models suggest that predictable, chronic mild stress levels in adolescence can actually increase resilience in adulthood [27], research suggests that many adolescents are coping with more severe stress levels. For example, in one study, 22% of adolescents aged 15-17 years reported moderate to severe levels of perceived stress [28]. In addition, a national survey conducted by the American Psychological Association (APA) in 2013 found that adolescents reported higher levels of stress than they perceived to be healthy and that their stress levels during the school year exceeded those of adults [29]. In addition, 31% of these adolescents reported that their stress levels had increased over the previous year, whereas 34% believed that these would increase over the following year [29]. These high stress levels may reflect a variety of stressors in adolescence. Qualitative interviews with adolescents suggest that they perceive their stress to stem from family and academic sources as well as role transitions and societal problems [30]. Quantitative research corroborates these accounts, suggesting that adolescents’ primary sources of stress are school and arguments at home [31]. In fact, adolescents report higher levels of stress during the school year than during the summer, with more than twice as many adolescents reporting extreme levels of stress during the school year than during the summer break [29]. Regardless of the source, such high levels of stress can have deleterious effects on adolescents’ psychological and physical health, including lower life satisfaction [32], poorer academic

https://www.researchprotocols.org/2021/4/e25545

JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e25545 | p.40

(page number not for citation purposes)
performance [33,34], cigarette smoking [35], emotional eating [36], poorer diet [37], and more frequent subjective health complaints (eg, headaches, fatigue, and sleep difficulties) [38]. Stress in adolescence has also been linked to internalizing symptoms, including depression [23,24] and anxiety [24]. Interpersonal stress, in particular, predicts the onset of the first major depressive episode among adolescents [39]. Thus, addressing stress in adolescents is critical for avoiding long-term problems. Although the frequency of stressful or negative life events predict negative outcomes [35,39], adolescents’ cognitive appraisal of these events, that is, their perceptions of stress, also predicts negative outcomes [32,36-38]. Perceived stress may be more strongly tied to negative outcomes than to life events alone [40].

However, adolescents tend not to believe that stress negatively affects their mental health, and most do not engage in regular stress management [29]. Approximately 55% of adolescents report setting aside time for stress management only “a few times a month or less,” 13% report never setting aside time for stress management, and only 5% report having seen a mental health professional about their stress [29]. On the basis of these data, the APA noted the need for opportunities to help adolescents address and cope with their stress “to break this unhealthy legacy of stress in America” [29].

Although there has been an increase in interventions targeting perceived stress in adolescence [41], they are much less common than those for anxiety and depression [42]. Furthermore, most interventions for stress are school-based, and the evidence for these interventions is mixed. In a systematic review of stress management interventions, only 58% of the reviewed studies found significant improvements in physiological indicators of stress (eg, blood pressure) or self-reported stress [43]. More recent meta-analyses of school-based programs suggest that the benefits may be limited to targeted samples (eg, with high levels of stress) and not universal programs [41,42] and may only be effective for school rather than social stress [41]. Given these findings, we need more research exploring the effectiveness of stress management programs, particularly those that could be delivered outside schools.

**Negative Cognitions and Stress**

Research suggests that stress in adolescence stems, at least in part, from feelings of helplessness and negative affect [44]. Therefore, effective interventions for adolescent stress may need to target the underlying negative cognitions as well as perceived stress. However, recommended approaches to addressing stress tend to focus on stress management training, relaxation training, and problem-solving and decision-making skills training [16,41,42], which may not adequately address the related negative cognitions.

Indeed, there is evidence that improving the content of adolescents’ cognition can help mitigate the negative effects of stress. For example, interpersonal stress predicts depressive symptoms in adolescents only when coupled with negative cognitions [39] or persistent low positive affect [45], whereas self-compassion buffers against the negative effects of perceived stress on internalizing symptoms [24]. In other words, addressing both negative cognitions and stress together may be particularly important in mitigating the negative consequences of stress [39].

The cognitive processes related to these negative cognitions may also be relevant. Rumination is a pattern of thinking in which individuals repeatedly think about the causes, consequences, and symptoms of their negative affect and often occurs following stressful events [46]. Rumination can be beneficial, where the individual engages in reflective pondering, or harmful, where the individual engages in brooding or moody pondering [47]. Similar to stress, rumination is another important transdiagnostic risk factor in prevention interventions for adolescents, as it predicts a host of other problems [48,49]. In adolescents, rumination, and particularly brooding, predicts depressive symptoms [50,51], anxiety [52], executive functioning impairments in selective attention and attentional switching [53], substance use problems [54], posttraumatic stress disorder following traumatic events such as terrorist attacks [55], and less sleep [48]. Furthermore, rumination and psychological distress appear to be involved in a cyclical pattern whereby ruminate thinking leads to psychological distress, which then predicts more rumination [48].

As rumination can occur in response to stressful events, it also plays a role in the negative consequences of chronic stress. For instance, longitudinal research has shown that more frequent stressful life events predict increased rumination among both adults and adolescents and that rumination mediates the relationship between stressful life events and anxiety symptoms in adolescents [56]. Other studies have shown that adolescents who ruminate in response to stress are at a greater risk of depression and substance misuse [57]. In college students, stress-reactive rumination, but not momentary ruminative self-focus, predicted depressive symptoms, although both predicted depressive symptoms among students with higher levels of stress [58].

In addition to contributing to the negative effects of stress, rumination may increase perceived stress. In one study, rumination exacerbated the relationship between life hassles and depression, anxiety, and stress in adolescents [59]. Rumination also increases adolescents’ likelihood of experiencing interpersonal stress, which in turn predicts more internalizing symptoms [60]. Therefore, applying more traditional approaches such as cognitive restructuring, self-monitoring, acceptance strategies, and attention control training [16] to address negative cognitive processes such as rumination may be helpful in reducing stress as well.

Unfortunately, aside from studies specifically testing stress management interventions, few intervention studies have included stress as an outcome variable [42]. To the best of our knowledge, no intervention research has focused on stress and rumination together. However, research suggests that similar approaches might be effective for reducing both stress and rumination. For example, researchers posit that mindfulness helps prevent depression by promoting low levels of rumination and high levels of self-compassion [61]. Indeed, trait mindfulness mitigates the relationship between life hassles and depression, anxiety, and stress [59], and mindfulness-based interventions have been shown to reduce perceived stress and
increase self-compassion in middle school students [62]. Cross-sectional research with college students also suggests that self-compassion attenuates the effects of rumination on stress [63], and increases in self-compassion following intensive meditation retreats predict improvements in perceived stress, rumination, negative affect, and depressive symptoms [64]. CBT has also been used successfully for both rumination [49,65] and perceived stress [66], although both outcomes have never been measured together. These findings suggest that there may be a common pathway for reducing both rumination and perceived stress in adolescents. Therefore, identifying interventions that can address both would be beneficial as they would target two, rather than one, important transdiagnostic risk factors in this population.

**Objectives**

The aim of this study is to test the effects of a digital mental health program that draws on several approaches to address perceived stress and rumination in adolescents. Happify for Teens is a version of the Happify Health platform that has been modified for those aged between 13 and 17 years. As in the adult version, Happify for Teens consists of gamified versions of evidence-based activities adapted from CBT [67], mindfulness-based stress reduction [68], positive psychology [69-71], behavioral activation [72,73], acceptance and commitment therapy [74], and psychoeducation [75]. Thus, Happify Health programs integrate a variety of recommended approaches to address negative cognitions [16] that appear to have promising effects for reducing perceived stress as well [62,66,76,77].

Although Happify for Teens has yet to be tested empirically, research on the adult version demonstrates that Happify Health programs lead to significant improvements in mental health. Observational studies of existing Happify Health users found significant improvements in subjective well-being after 6-8 weeks of use, with greater gains among users who completed more activities [78-80]. Similarly, participants in randomized controlled trials (RCTs) who completed at least two activities per week had significant improvements in depressive symptoms, anxiety, and resilience compared with participants in a psychoeducation control group or participants who completed less than 2 Happify Health activities per week [81,82].

Notably, although published analyses did not examine the effect of Happify Health on perceived stress directly, perceived stress was included as 1 of 3 components for a resilience index, which significantly decreased in users who competed at least two activities per week for 8 weeks [81,82]; unpublished data indicate that perceived stress specifically decreased along with resilience (Parks, AC, unpublished data, December 2018). In addition, experimental research suggests that using one of the activities within the Happify Health platform, a heart rate variability biofeedback game, following a stressful event may help users manage their stress, as evidenced by lower salivary alpha amylase in participants who completed this activity relative to controls [83]. These data suggest that Happify for Teens may help adolescents manage their stress while providing the added benefit of addressing underlying negative cognitions and cognitive processes.

To test the effect of Happify for Teens on perceived stress in adolescents, we plan to recruit adolescents aged 13-17 years who report elevated levels of perceived stress and rumination. Participants will be randomly assigned to complete 8 weeks of activities via Happify for Teens or to a corresponding wait-list control group. We will then compare changes in perceived stress at 4 weeks, 8 weeks, and 1 month post intervention across the 2 groups. We will also examine changes in brooding (the maladaptive component of rumination) as a secondary outcome. Finally, as sleep disturbance [84,85], optimism [86,87], and loneliness [88,89] have been identified as other potential transdiagnostic risk factors for future mental health impairment that may begin in adolescence, we also examined changes in sleep disturbance, optimism, and loneliness as additional exploratory outcomes.

**Methods**

**Participants**

This study is an RCT (NCT04567888) with a target sample size of 800 participants (400 participants per condition). Power analyses indicated that 200 participants (100 per condition) would be sufficient for 80% power to detect a small effect; however, attrition rates for this study are difficult to predict. Previous RCTs using Happify Health with adult participants had response rates ranging from 56% to 72% for an 8-week intervention and posttest [81,82]. Research on other digital interventions with adolescents has reported substantial variability in completion rates. The levels of noncompletion for internet-based CBT interventions range from 33.3% to 69.6%, with low use rates overall [90], whereas for digital interventions targeting diet and physical activity, completion rates range from 37% to 100% [91]. On the basis of this variability, we opted for a conservative approach to ensure adequate power and estimated a 75.0% (600/800) attrition rate, leading to an initial target sample size of 800 adolescents to maintain adequate power.

**Recruitment**

To recruit participants, we will advertise the study to existing Happify Health users who indicate that they have adolescent children. In addition, to target parents (and adolescents) less familiar with the platform, we will also advertise the study using targeted advertising on social media. Finally, we will distribute information about the study in 2 public schools (one located in California and another in Montana) that contacted Happify Health about the program. Representatives from these schools who contacted our research staff to express interest in the product will distribute letters to parents within their schools, with information about the study and a link to the parent prescreening survey.

Regardless of the recruitment method, interested parents will be directed to complete a brief screening questionnaire via Qualtrics to verify that their child is aged between 13 and 17 years, resides in the United States, and has never used Happify Health before. After meeting the initial eligibility criteria, parents will provide consent for each eligible child to participate in the study. In addition, we plan to use a snowballing recruitment method whereby parents who complete this...
screening questionnaire will be encouraged to share the study with other parents, and they will be entered into a drawing for a US $50 Amazon gift card for each parent they refer to the screening questionnaire.

We will then send eligible adolescents an email invitation to complete a separate screening questionnaire via Qualtrics, consisting of questions regarding their age, country of residence, previous Happify Health use, and about the study procedures to confirm their eligibility for the study. Respondents will also complete the Perceived Stress Scale (PSS) [40] and Ruminative Response Scale (RRS)—Short Form–Brooding Subscale [47]. Adolescents will be eligible to participate if they are aged between 13 and 17 years, currently reside in the United States, have never used Happify Health, report elevated levels of stress (ie, PSS score >14) and rumination (ie, RRS–Brooding Subscale score ≥10), and indicate that they are willing and able to complete the study activities. Participants from the same household will be permitted to participate in the study but will be instructed not to discuss the study or the platform with each other.

Compensation
Participants will be compensated with US $20 for completing each of the 4-week, 8-week, and 1-month postintervention assessments. Participants who complete all 4 assessments (baseline and 3 other assessments) will receive a US $20 bonus. Thus, participants will receive up to US $80 as compensation for their time; participants will be compensated in the form of an electronic gift card.

Primary Outcome Measure: PSS
The PSS [40] is a widely used measure of the extent to which respondents view their lives as unpredictable, uncontrollable, or overloaded and has been used with adolescents in other studies [92,93]. It consists of 10 items asking participants the extent to which they have felt each of the feelings and thoughts in the previous month (eg, “In the last month, how often have you felt nervous and ‘stressed’?”), and items are rated on a scale from 0 (never) to 4 (very often). Ratings are summed so that higher scores indicate greater perceived stress.

Secondary Outcome Measure: RRS–Short Form–Brooding Subscale
The brooding subscale of the RRS [47] examines the extent to which respondents engage in moody pondering and has been validated with adolescents [94,95]. Participants indicate how often they engage in each of the 5 behaviors (eg, “Think ‘Why can’t I handle things better?’”) on a scale from 1 (almost never) to 4 (almost always). For this study, instructions were modified so that participants can indicate the extent to which they engaged in these behaviors during the previous month. Ratings are summed so that higher scores indicate that participants engaged in more brooding during that time.

Exploratory Outcome Measures
Life Orientation Test–Revised
The Life Orientation Test–Revised is a 10-item measure of optimism [96] that has been used with adolescents [97]. Participants indicate the extent to which they agree with each statement (eg, “In uncertain times, I usually expect the best”) on a 5-point scale, ranging from 0 (strongly disagree) to 4 (strongly agree). A total of 4 items are filler items, and ratings on the remaining 6 items can be summed to obtain an overall optimism score so that higher scores indicate more optimism.

Patient-Reported Outcomes Measurement Information System Pediatric Sleep Disturbance Scale–Short Form 4a
This is a 4-item scale measuring the extent to which participants experienced sleep disturbances over the past 7 days (eg, “In the past 7 days, I had trouble sleeping”), appropriate for respondents aged between 8 and 17 years [98]. Each item is rated on a scale from 1 (never) to 5 (always), and ratings are summed so that higher scores indicate more sleep disturbance or poorer sleep quality.

Roberts UCLA Loneliness Scale-8
The Roberts UCLA Loneliness Scale-8 [99] was adapted from the original 20-item UCLA loneliness scale [100] for use with adolescents. It consists of 8 items (eg, “I lack companionship”), and participants indicate how often each of the statements are descriptive of them on a 4-point scale from never (0) to often (3). Ratings are summed so that higher scores indicate higher levels of loneliness.

Usage Statistics
In addition to self-reported outcomes, we also plan to collect data on usage and engagement via the Happify Health platform for those assigned to the intervention condition. Specifically, we will passively collect information about the activities completed and the number of active days.

Procedure
This study will take place over approximately 3 months, including an 8-week intervention period and a 1-month follow-up. See Table 1 for the schedule of activities. All study procedures were reviewed and approved by IntegReview, an independent institutional review board.

After obtaining parental consent, eligible participants will be contacted via email with instructions to complete an electronic assent form and the baseline assessment. This assessment will include all primary, secondary, and exploratory outcome measures. To assess data quality, we plan to include an attention check within each questionnaire in the baseline assessment and subsequent assessments. These attention checks will instruct participants to select a specific response option to determine whether participants are reading items carefully (eg, for the PSS, an attention check item could instruct the participant to select Never).

Once participants have completed the baseline assessment, they will be randomly assigned to either the intervention group or the wait-list control group and will receive further instructions via email. Participants who do not complete the baseline assessment will not be assigned to a condition and will be disqualified from the study. Similarly, participants who fail all
attention checks within the baseline assessment will not be assigned to a condition and will be disqualified from the study. Participants in both conditions will receive emails or text messages throughout the study to check in and keep them engaged. At 4, 8, and 12 weeks (1 month post intervention), all participants will also be prompted to complete all primary, secondary, and exploratory measures via Qualtrics. At each assessment, participants will also be asked to indicate whether they participated in any other interventions or used any digital self-help or wellness program since the last assessment date. Participants will receive reminders if they have not completed these assessments on time, and if a participant has not completed an assessment within 7 days of its scheduled date, a member of the research team will reach out via text message or email to check in.

<p>| Table 1. Schedule of activities for prescreen, intervention period, and follow-up assessments. |</p>
<table>
<thead>
<tr>
<th>Assessments</th>
<th>Time</th>
<th>Prescreen</th>
<th>Baseline</th>
<th>Week 4: mid-point assessment</th>
<th>Week 8: postintervention assessment</th>
<th>Week 12: 1-month follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Stress Scale</td>
<td>✓ ✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Secondary outcome</strong></td>
<td></td>
<td>✓ ✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruminative Responses Scale–Short Form–Brooding Subscale</td>
<td>✓ ✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Exploratory outcomes</strong></td>
<td></td>
<td></td>
<td>✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Orientation Test–Revised</td>
<td>_ _ _ b</td>
<td>✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient-Reported Outcomes Measurement System Pediatric Sleep Disturbance Scale–Short Form 4a</td>
<td>—</td>
<td>✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberts UCLA Loneliness Scale</td>
<td>—</td>
<td>✓ ✓ ✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Indicates that the outcome was assessed at this time.

*b* Indicates that the outcome was not assessed at this time.

**Intervention Group: Happify Health for Teens**

Participants assigned to the active intervention will receive instructions to download Happify Health and then be directed into the Teens platform after creating an account. The Happify for Teens platform was developed using the same model as the original Happify Health platform; intervention activities are organized into 6 different skills: *savor* (mindfulness skills), *thank* (gratitude), *aspire* (optimism, goal-setting, and finding meaning or purpose), *give* (kindness, forgiveness, and prosocial behavior), *empathize* (self-compassion and perspective-taking), and *revive* (physical health). However, for the adolescent platform, these activities were modified and then reviewed by a panel of adolescents to ensure that the language and content were relevant to that population (*Figure 1*).

**Figure 1.** Screenshot of Happify for Teens.

Activities from different categories are then organized into *tracks* focused on addressing a specific area of concern, such as increasing confidence or reducing negative thinking (see *Figure 2* for a sample track description). Users can select a track that interests them, and within each track, they can choose what activities they prefer to complete, thereby offering users a more personalized experience in terms of selecting tracks as well as activities, which has been shown to improve efficacy [101,102].
Although participants will have access to all Happify for Teen tracks, we will feature a track called Stop the Worry Cycle. Each track consists of 4 different parts, and users must complete activities within each part to earn a silver or gold medal before they can move on to the next part. Users can also complete activities on demand through instant play and can switch tracks at any time.

Participants will not be given explicit instructions on how often they should use the platform or how many activities to complete; however, we will encourage them to engage with the platform daily. Participants will also receive push notifications on their mobile device every other day to remind them to access the platform, and they will receive weekly emails as part of the Happify Health platform to help increase engagement. If a participant has not completed any activities within the platform for 7 days, they will also receive an email or text message (based on participant preferences) from a member of the research team to check in. After the 8-week intervention period, participants in the intervention group will continue to have access to the Happify for Teens platform until the 12-week assessment but will receive no explicit instructions or no push notifications to use the program following the intervention period. We will also no longer track use to contact participants who have not used the program during this 1-month period.

**Figure 2.** Screenshot of track description for “Stress Buster 101” track in Happify for Teens.

**Wait-list Control Group**

Active intervention or Happify Health for Teens will be compared with a wait-list control condition. Participants assigned to this condition will be instructed via email that they have been assigned to a wait-list control and that they will be given access to the Happify for Teens platform after the 12-week study period; however, they are still expected to complete the study assessments.
Data Analysis

Changes in Perceived Stress and Secondary or Exploratory Outcomes

To assess whether changes in participants’ perceived stress differed across the 2 groups, we plan to conduct a 2 (group: intervention vs control) × 4 (time: baseline, 4 weeks, 8 weeks, and 12 weeks) repeated-measures ANOVA (Analysis of Variance) on participants’ PSS scores, controlling for participant age, gender, and race. Similar analyses will be conducted on the secondary and exploratory outcome measures. In addition, to determine whether usage predicted changes in these outcomes, we will also conduct regression analyses among participants in the intervention condition, regressing their scores on each outcome variable on the number of activities completed within the Happify for Teens platform, while controlling for corresponding baseline scores on that outcome.

Rumination as a Potential Mediator

To test whether any observed changes in perceived stress can be attributed to changes in brooding, we also plan to conduct mediation analyses using Hayes PROCESS macro for SPSS [103].

Data Exclusion

As careless responding is problematic with web-based surveys and may artificially increase relationships among variables [104], we plan to use 2 separate a priori mechanisms to identify low-quality data: (1) failing 3 or more attention checks in an assessment and (2) completing an assessment at a rate faster than 1 second per item [104]. We will then rerun the analyses described above without these participants to determine whether any effects differ after removing low-quality data.

Results

Recruitment is expected to begin during the second quarter of 2021 and will continue until 800 participants have been randomized to one of the 2 conditions. Participants will begin the 8-week intervention after being screened and consenting to participate; therefore, the first wave of data collection is expected 8 weeks after recruitment begins. We estimate that all participants will complete the intervention in the third quarter of 2021, with final completion, including the 1-month follow-up, by the end of the third quarter of 2021. Given uncertainty regarding attrition and use in an adolescent population, if fewer than 200 participants complete the postintervention assessment, we plan to conduct a second round of recruitment to increase our sample size to ensure adequate power.

Discussion

Recent data indicate that adolescents are experiencing increasing levels of stress, particularly in school, when their stress levels exceed those of adults [29]. Given that chronic stress predicts a host of psychological, behavioral, and physical problems in adolescence [28,32-38], developing opportunities to help adolescents cope with their stress has become increasingly important [29]. Several interventions for adolescent stress have been developed; however, the evidence supporting their effectiveness is mixed [41-43]. Many of these interventions rely on traditional approaches to treating stress, including stress management training, relaxation training, and problem-solving and decision-making skills training [41,42]; however, these approaches may not adequately address negative cognitions that contribute to perceived stress [44]. For example, rumination, or maladaptive patterns of moody pondering, contributes to perceived stress [59,60] and exacerbates the negative consequences of stress [57,58]. Therefore, interventions may be particularly effective when they also incorporate approaches to addressing negative cognitions, such as cognitive restructuring and acceptance strategies [16]. To our knowledge, this is the first study to examine the effects of a stress management program on both perceived stress and rumination in adolescence and to explore whether changes in rumination mediate the effects of the program on stress.

Evaluating the Efficacy of a Digital Prevention Intervention

Currently, most stress management programs for adolescents continue to be delivered within schools [41-43]. Although school-based prevention programs can be effective [42,43], there are numerous barriers to implementing evidence-based interventions within schools, including lack of time and resources and financial constraints [105]. Costs associated with implementing such interventions may be particularly prohibitive to schools in socially and economically disadvantaged areas [42], where students may need these interventions the most [41]. Targeted mental health interventions may have additional barriers to student participation due to fear of stigmatization [106]. Given that 95% of American adolescents own or have access to a cellular phone and 88% have daily access to a computer [107], a digital intervention addressing stress could offer a better, and arguably more cost-effective, opportunity to reach more adolescents. In addition, participation in the program would not require an entire school to buy-in; therefore, the decision to participate would be individual and private, potentially reducing concerns with stigma as well.

Although mobile mental health apps have become increasingly popular [108], mobile apps developed specifically for youth are still relatively scarce [109]. Furthermore, despite the popularity of mobile mental health apps, high quality empirical research on these apps in general is lacking [109-111]. For example, in a review of mobile mental health apps listed in the National Health Service app library, approximately 28% of mobile apps for managing depression and anxiety provided evidence to support their effectiveness claims, and only 14% of apps used validated clinical outcomes to do so [112]. Moreover, the market for these mobile apps is volatile, with one depression-focused app disappearing from the marketplace every 3 days, on average [113]. Early and robust tests of these platforms are therefore an important and necessary step to legitimize digital interventions and mobile mental health apps [114].

Empirical tests of digital interventions are particularly important with regard to platforms developed specifically for youth, as it is less clear how adolescents will respond to these interventions.
Some research suggests that adolescents prefer therapy when delivered digitally rather than face-to-face [115]. However, other research suggests that adolescents report high internet and mobile app use in general but low use of mental health apps specifically [116]. Some adolescents also report feelings of being labeled and stigmatized if using a digital mental health intervention, even if no one is aware that they are using one [117]. In particular, when considering stress, adolescents tend to believe that stress has little effect on their mental health and rarely engage in stress management on their own [29]. Many studies testing digital interventions in adolescence have aimed to address poor engagement by including some degree of therapist guidance [118], phone contact [119], and physician involvement [120]; however, the effects of these strategies are inconclusive [90]. Consequently, although several studies have demonstrated the effectiveness of the general Happify Health platform with adults [78-81,92], the extent to which adolescents will engage with the digital platform, and in turn, the extent to which they will benefit from the program, is difficult to predict.

Strengths and Limitations

Research on adolescent interventions has been characterized by low power. Although some studies of school-based stress management interventions have had large samples, a meta-analysis of 54 studies reported that 52% of the included studies had fewer than 100 participants and 72% had fewer than 200 participants [41]. Studies on digital interventions have comparably smaller samples. For example, meta-analyses of studies on digital interventions and mobile mental health apps reported sample sizes ranging from 2 to 206 participants [90,109]. Given the mixed findings in previous research, particularly with stress management interventions [41,42], studies with larger sample sizes are needed to clarify the effectiveness of these interventions. Considering the high dropout rates with digital interventions in general [81,82,121-123], and specifically with adolescents [90,91], this study was designed to be adequately powered even with 75% attrition.

Another limitation of previous research is that the focus on school-based interventions likely results in samples that are geographically bound to schools in which these programs can be delivered, which may reduce the generalizability of those results. By comparison, testing a digital program that is accessible on adolescents’ personal computers or smartphones eliminates many barriers of school-based research, which may result in a more diverse and representative sample of adolescents.

However, our participants may differ from the general population in other ways. First, as we are advertising the study to existing Happify Health users and within specific schools that expressed an interest in the Happify Health platform, our sample may not be representative of the general adolescent population in the United States. Although this may limit the generalizability of our findings to the general population, our recruitment strategy is likely to yield a sample that is representative of Happify Health users, and consequently, our findings should be more representative of our population of interest: adolescents who will be interested in using the Happify Health platform.

In addition, with the goal of increasing retention and use, we require participants to complete the baseline assessment before qualifying for the study, and we plan to contact participants when they have not engaged with the platform or have not completed an assessment for 7 days. Consequently, our sample may include adolescents who are particularly motivated and conscientious. In addition, because we are advertising to caregivers, adolescents who participate in the study may also differ from the general population in terms of attachment style, quality of relationship with their caregiver, parenting styles, or even the caregiver’s own mental health. All these factors could affect the extent to which adolescents respond to the program [124,125].

Our sample also represents adolescents who report elevated levels of both perceived stress and brooding; therefore, it will remain unclear whether Happify for Teens would be effective as a universal stress management program (ie, for adolescents in general). Similarly, whether Happify for Teens would be effective for adolescents with either elevated levels of perceived stress or rumination, rather than elevated levels of both, will also remain unclear. Although this limits generalizability, targeted programs tend to have stronger effects than universal programs [126], and school-based stress management programs appear to be effective only with selective samples [41]. Therefore, we felt it was important to first examine whether Happify for Teens is effective as a targeted program. Arguably, given that adolescents with higher levels of perceived stress and rumination also have a higher risk of developing more severe impairment without intervention [23,24,50,52], a targeted approach also helps to test efficacy more robustly in the population that needs the program the most. Nevertheless, future research should test the effect of this program with a less selective sample of adolescents.

Previous research on digital adolescent mental health interventions has also been criticized for a lack of follow-up assessments [90,127,128]; however, some interventions, including school-based stress management interventions, may have stronger effects over time than immediately post intervention [41], emphasizing the importance of long-term follow-up. This study includes a 1-month follow-up, providing some insight into the longitudinal effects of the program, but future research should include more long-term follow-up to better ascertain whether effects diminish or strengthen over time. In addition, although reducing perceived stress and rumination should reduce participants’ likelihood of developing mental health problems, such as depression and anxiety, further research is needed to assess the long-term implications of Happify for Teens on other mental or physical health outcomes.

Conclusions

Adolescent mental health is a growing area of concern, as many mental health problems begin in adolescence and continue into adulthood [5,6]. Prevention interventions offer a means of promoting mental health among youth before more serious dysfunction or impairment becomes an issue, thereby reducing the likelihood of mental health disorders and associated chronic
disorders [2,13]. Such interventions are not only effective [2,17-22] but may also help reduce the burden on mental health services from adolescents with little to no impairment [13], who represent an increasing proportion of adolescents seeking services [9]. These interventions are most effective when they target samples with a higher risk of mental health problems [2,41] and address transdiagnostic risk factors [15,16]. Stress is an important transdiagnostic risk factor that has been studied primarily in the context of school-based stress management programs [41-43] and without considering underlying negative cognitions that might contribute to perceived stress and exacerbate the negative effects of that stress [39,44,57-60]. This trial tests a novel digital stress management program for adolescents and its effects on perceived stress and rumination. These data will, therefore, provide important information about the potential efficacy of a more scalable and cost-effective method of improving perceived stress and other associated outcomes in an adolescent population.

Authors’ Contributions

EMB contributed to designing the study and wrote the initial draft of this protocol. HEW contributed to designing the study and to writing the protocol. JLS contributed to designing the study and provided general manuscript support for this protocol. ACP contributed substantially to the study aims, scope, and design of the study.

Conflicts of Interest

EMB, HEW, and ACP are employees of Happify Health. JLS was employed by Happify Health when the study was designed and this protocol was written.

References


83. Hunter JF, Olah MS, Williams AL, Parks AC, Pressman SD. Effect of brief biofeedback via a smartphone app on stress recovery: randomized experimental study. JMIR Serious Games 2019 Nov 26;7(4) [FREE Full text] [doi: 10.2196/15974] [Medline: 31769761]


107. Smartphone access nearly ubiquitous among teens, while having a home computer varies by income. Pew Research Internet on Twitter. URL: https://www.twitter.com/pin/362759449480087440 [accessed 2020-10-29]


121. Donkin L, Christensen H. Smartphone access nearly ubiquitous among teens, while having a home computer varies by income. Pew Research Internet on Twitter. URL: https://www.twitter.com/pin/362759449480087440 [accessed 2020-10-29]

122. Boucher et alJMIR RESEARCH PROTOCOLS


A Web-Based Lifestyle Intervention Aimed at Improving Cognition in Patients With Cancer Returning to Work in an Outpatient Setting: Protocol for a Randomized Controlled Trial

A Josephine Drijver\textsuperscript{1,2}, MSc; Jaap C Reijneveld\textsuperscript{1,2,3}, MD, PhD; Linda M P Wesselman\textsuperscript{2,4,5}, PhD; Martin Klein\textsuperscript{2,6}, PhD

\textsuperscript{1}Department of Neurology, Amsterdam UMC, Amsterdam, Netherlands
\textsuperscript{2}Vrije Universiteit Amsterdam, Amsterdam, Netherlands
\textsuperscript{3}Stichting Epilepsie Instellingen Nederland (SEIN), Heemstede, Netherlands
\textsuperscript{4}Alzheimer Center Amsterdam, Department of Neurology, Amsterdam UMC, Amsterdam, Netherlands
\textsuperscript{5}Amsterdam Neuroscience, Amsterdam, Netherlands
\textsuperscript{6}Department of Medical Psychology, Amsterdam UMC, Amsterdam, Netherlands

Corresponding Author:
A Josephine Drijver, MSc
Department of Neurology
Amsterdam UMC
ZH 2 A 63
De Boelelaan 1117
Amsterdam, 1081 HV
Netherlands
Phone: 31 20444 4219
Email: a.drijver@amsterdamumc.nl

Abstract

Background: A high percentage of patients with cancer experience cognitive impairment after cancer treatment, resulting in a decreased health-related quality of life and difficulty returning to work. Consequently, there is a need for effective treatment options to improve cognitive functioning in these patients. In a healthy aging population, multidomain web-based lifestyle interventions have been found to be effective in preventing cognitive decline and improving cognitive functioning.

Objective: This study aims to investigate the feasibility and effectiveness of the web-based lifestyle intervention Mijn Fitte Brein (My Fit Brain [MFB]) on cognitive functioning in patients with cancer returning to work.

Methods: The study consists of a feasibility study (N=10), followed by a randomized controlled trial (RCT; N=220). Patients will be recruited by their occupational physicians after their return to work following cancer treatment. Mijn Fitte Brein is organized into 4-week cycles in which patients set a lifestyle goal using the Goal Attainment Scale, receive weekly tips and support, and finally evaluate whether they succeeded in achieving this goal. Lifestyle goals are based on 6 domains: physical exercise, diet, sleep, stress, alcohol use, and smoking. In the feasibility study, data on user experience (structured interview) and usability, assessed with the Post-Study System Usability Scale, will be collected and used to optimize Mijn Fitte Brein. In the RCT, patients will be randomized 1:1 between an intervention group and a control group. Patients will be assessed at baseline, 3 months, and 6 months. The primary outcome measure is subjective cognitive functioning, assessed with the Functional Assessment of Cancer Therapy–Cognitive Function (FACT-Cog). Secondary outcome measures are lifestyle, objective cognitive functioning, and work and psychosocial factors.

Results: Recruitment for the feasibility study has started in February 2020. As of July 2020, however, no patients have been enrolled (due to COVID-19 restrictions). The findings of the feasibility study will be used to optimize the Mijn Fitte Brein intervention. Enrollment for the RCT will continue when possible. The feasibility study will take 6 months (including making adjustments to the intervention), and the RCT will take 2 years. The final results are expected in 2024. The results of the feasibility study and the RCT will be published in peer-reviewed journals.

Conclusions: This is the first time the feasibility and efficacy of a multidomain web-based lifestyle intervention will be studied in patients with cancer. If Mijn Fitte Brein is found to be effective in decreasing cognitive complaints in these patients returning to work, it will be a promising treatment option because of being both affordable and accessible.
Introduction

Cancer and cancer-related treatment can have detrimental effects on cognitive functioning, and complaints of cancer-related cognitive impairment are indeed reported by up to 75% of patients [1]. Besides these cognitive complaints, a subgroup of patients also shows objective cognitive impairment on neuropsychological tests [2]. These findings are of concern, because cognitive impairment negatively affects health-related quality of life (HRQoL) in patients with cancer and survivors of cancer [3]. Not all patients with (mild) cognitive impairment after cancer treatment have cognitive complaints and not all patients who have cognitive complaints show impairment on neuropsychological tests. Therefore, both subjective and objective measures are of interest when studying cognition in this patient group. In clinical practice both outcome measures are useful for patient education and counseling.

In patients with cancer, cognitive impairment affects the ability to work, which also impacts HRQoL [4]. Patients report not being able to keep up with their previous workload and experience a lack of understanding from their environment [5]. Cognitive problems are more frequently reported in patients who experience persistently low work functioning after return to work [6]. Because of these issues, patients are often not able to return to their previous roles or are not able to return to work at all. This highlights the importance of finding treatment options aimed at improving cognitive functioning in patients with cancer returning to work.

Besides the type of antitumor treatment, multiple predictors of cognitive impairment have been identified in patients with cancer [7]. Most of these predictors, such as genetic and sociodemographic factors, are nonmodifiable. Lifestyle, however, is a modifiable predictor and therefore interesting as a target for treatment. Multiple lifestyle factors such as low physical activity, impaired sleep, and unhealthy diet have been identified as risk factors for cognitive decline in patients with cancer [8-11]. Research in healthy individuals and several patient groups has shown that lifestyle interventions are able to improve cognitive functioning or prevent cognitive impairment [12]. In the field of cancer, while several lifestyle interventions have been shown to improve cognitive functioning, most interventions focused on a single lifestyle domain, most frequently targeting physical activity. Yoga, Tai Chi and Qigong, and aerobic exercises have been found to correlate with improved self-perceived cognitive functioning in patients with cancer [13-17]. Research in other populations has shown that interventions targeting multiple lifestyle domains are more effective than single-domain interventions. For example, the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) trial, which focused on diet, exercise, cognitive training, and vascular risk monitoring [18], was found to be effective in improving or maintaining cognitive function in an at-risk, but otherwise healthy, elderly population [19].

One disadvantage of these face-to-face multidomain lifestyle interventions such as the FINGER trial is that multiple weekly visits to the training and counseling facility are required [20]. Besides the burden of multiple visits, the costs of providing/offering these interventions are considerable. More affordable alternatives are web-based interventions, which have been studied in diverse populations [21]. Such interventions are not only more affordable, but are also easily accessible for patients, as currently 97% of the Dutch population has internet access [22]. Recently the Healthy Ageing Through Internet Counselling in the Elderly (HATICE) trial studied the effect of an internet-supported intervention aimed at decreasing cardiovascular risk factors and the risk of cognitive decline in an older population [23]. The first results showed that this web-based intervention was feasible and effective [24].

In patients with cancer, no web-based multidomain lifestyle interventions aimed at improving cognitive functioning have been studied so far. This protocol describes a feasibility study and a subsequent randomized controlled trial (RCT) aimed at evaluating the feasibility and effectiveness of such an intervention in patients with cancer returning to work. In a self-motivated intervention, patients do not receive guidance from a counselor and are expected to work through the program independently. The study aims to establish the feasibility of a self-motivated lifestyle intervention to this population, and whether the program is beneficial for the cognitive functioning of patients. In addition, the underlying factors mediating effectivity will be studied.

Methods

Approval

The study protocol has been approved by the Medical Ethical Committee of the VU Medical Center. The study will be conducted according to the principles of the Declaration of Helsinki, and in accordance with the Medical Research Involving Human Subjects Act (WMO). The trial has been registered at the Netherlands Trial Registry (NL8407).

Design

Baseline Assessment

This study consists of a feasibility study and an RCT (Figure 1) in an outpatient setting. Participants will be informed of the study by their physician at a Dutch occupational health institute.
Baseline assessment will take place at the hospital or at the participant’s home, depending on the participant’s preference. A member of the research team will answer any of the participant’s questions before informed consent is given. The researcher, who has been trained on the test protocol, administers a neuropsychological test battery and the participant fills in the questionnaires (see the “Outcome Measures” section) online. The first visit will take approximately 2 hours. Follow-up assessments will adhere to the same protocol.

The information on the RCT protocol is provided according to the CONSORT 2010 checklist (Multimedia Appendix 1) [25].

**Feasibility Study**

In the feasibility study, 10 participants will get access to the Mijn Fitte Brein (My Fit Brain [MFB]) intervention for 3 months.

After the baseline assessment the participants will receive an email with login information to access the web-based Mijn Fitte Brein intervention for 3 months. The participants are instructed to use Mijn Fitte Brein weekly during this period. At the 3-month follow-up, the cognitive assessment will be repeated and the participants will fill in the questionnaires. For this feasibility study specifically, the researcher will interview the participants about their experience with Mijn Fitte Brein. In addition, the participants will fill out a questionnaire about the usability of the platform.

**Randomized Controlled Trial**

In the RCT, 220 participants will be randomly assigned to the intervention group or the control group. The RCT uses a parallel design. Randomization will be performed by a researcher not associated with the testing of participants, in Castor EDC [26]. Variable block randomization will be used with blocks of 8, 10, and 12, with each condition (Mijn Fitte Brein/No treatment) randomly generated 50% within each block to ensure an unpredictable allocation sequence with equal numbers of participants in each group at the completion of each block. The allocation ratio is 1:1. Blinding of patients is not possible due to the nature of the intervention. Assessors of the neuropsychological examination will be blinded.

Participants in the treatment group will get a 6-month access to Mijn Fitte Brein, whereas those in the control group will get access after a 6-month waiting period. Participants will be assessed at baseline, 3 months, and 6 months. The participants will not be interviewed at the end of the study. Control group participants are allowed to find other help sources for their cognitive problems, and this is recorded in the study. However, they are not allowed to participate in other studies specifically aimed at improving cognition.

**Participants**

**Eligibility Criteria**

To be eligible to participate in this study, potential participants must meet all of the following criteria: be over 17 years of age, be employed, have received primary treatment (chemotherapy, immunotherapy) for a recent noncentral nervous system cancer, completed treatment (except for hormonal treatment), report cognitive complaints, and provide written informed consent. Exclusion criteria are primary or secondary central nervous system tumors, insufficient mastery of the Dutch language, self-reported insufficient computer skills, history of brain injury.
with loss of consciousness, history of brain surgery, currently under active treatment for psychiatric or neurodegenerative disorders, self-reported substance abuse, and severe visual impairments.

**Recruitment**

All participants will be informed of the study by their occupational health physician. If interested, participants will be approached by one of the researchers and receive further information. When inclusion criteria are met, the researchers will schedule the first visit. During this visit, all participants will provide written informed consent before the start of the study activities.

**Sample Size**

In the feasibility study 10 participants will be included. Regarding the subsequent RCT, thus far, no comparable studies have been published using a similar intervention in a population of cancer survivors. Therefore, sample size is calculated based on the minimal clinically important difference (MCID) on the Functional Assessment of Cancer Therapy–Cognitive Function (FACT-Cog; a questionnaire to measure subjective cognitive functioning in patients with cancer and the primary outcome of the RCT). The MCID was found to be 9.6 points when using an anchor-based approach [27]. In the same population, an SD of 23.3 was found. Using a significance level of $\alpha = 0.05$ and test power ($1 - \beta$) of 80% leads to a sample size of 188 participants equally divided into 2 groups. Assuming an estimated drop-out rate of 15%, a total of 220 patients will be included in the RCT.

**Intervention: Mijn Fitte Brein**

**Description**

Mijn Fitte Brein (translation: My Fit Brain) is a web-based lifestyle intervention based on the Brain Aging Monitor (BAM) [28]. Use of the BAM was associated with significant improvement in cognition-associated lifestyle factors in healthy older adults (40-60 years) [29]. The BAM has been adapted where necessary to conform to current lifestyle guidelines, to be suitable for the patients in this study, and to comply with the EU and European Economic Area General Data Protection Regulation (GDPR). Mijn Fitte Brein is built on the Minddistrict platform [30].

The Mijn Fitte Brein web-based intervention is constructed around 6 lifestyle domains: physical exercise, nutrition, sleep, stress, alcohol, and smoking. It consists of a lifestyle questionnaire followed by 3 modules (Figure 2): a goal setting module, an information and diaries module, and a goal evaluation module. See Multimedia Appendix 2 for the monthly schedule of Mijn Fitte Brein.

**Figure 2.** Screenshots showing the four steps of the MFB intervention. A: lifestyle questionnaire, B: goal setting, C: tips, D: goal evaluation, E: diary on the smartphone.
Questionnaire and Feedback
The first time a user logs onto the website, a questionnaire covering all 6 lifestyle domains is presented. Subsequently, the user receives feedback tailored to the provided answers. If national guidelines for the specific domains are available, the feedback states whether the user adheres to these guidelines. Otherwise, feedback includes a comparison of the user’s current behavior with the national average. It also contains general recommendations on how to improve lifestyle in the specific domains. Feedback concludes with recommendations about the domains where the participant could improve most. Finally, the user is asked to choose a domain to work with, and to set a goal for the upcoming month. See Table 1 for possible goals within each lifestyle domain.

Table 1. Domains and goals.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical exercise</td>
<td>Start moving, start exercising, increase exercising</td>
</tr>
<tr>
<td>Nutrition</td>
<td>Healthy eating(^a)</td>
</tr>
<tr>
<td>Smoking</td>
<td>Quit smoking, decrease smoking</td>
</tr>
<tr>
<td>Alcohol</td>
<td>Decrease daily alcohol consumption, decrease weekly alcohol consumption, decrease glasses of alcohol when drinking</td>
</tr>
<tr>
<td>Stress</td>
<td>Reduce stress</td>
</tr>
<tr>
<td>Sleep</td>
<td>Improve sleep quality, increase sleep duration, improve sleep pattern</td>
</tr>
</tbody>
</table>

\(^a\)For the goal healthy eating users can select 2 subgoals out of the following: “have breakfast more often,” “eat more vegetables,” “eat more fruit,” “eat more fish,” “eat less unhealthy snacks,” and “eat less often after dinner.”

Goal Setting
Next, the user is referred to the goal-setting module. First, the user receives an explanation of the goal-setting method that is used in Mijn Fitte Brein: Goal Attainment Scale (GAS) [31]. Second, following the GAS method, the user has to complete a scale ranging from –2 (I have made minimal progress) to 2 (I have done a lot better than my original goal). The user specifies his/her goal by describing a situation for each level on the scale.

For each potential goal, Mijn Fitte Brein provides instructions and recommendations that assist the user with realistic goal setting. To further aid goal setting, an example GAS and case description, as well as step-by-step instructions are presented. After the user has set a goal, he/she receives reinforcing feedback aimed at initiating behavior change. Last, the user specifies the goal by completing a personal GAS. For an example of the goal-setting module, see Multimedia Appendix 3.

Information and Diaries
The day after completing the goal-setting module, the user is given access to the first part of the training on the Minddistrict dashboard. Every week a new training module becomes available, providing tips and tricks tailored to the specific goal the user has chosen. For example, tips on how to overcome barriers in reaching your goal, mindfulness exercises, and training schedules for physical activity are shared. The user receives an email alert when a new part of the training becomes available.

In addition, daily diaries are made available for the user to track his/her progress. The user can access the diary on both a smartphone and a computer. The diaries are tailored to each goal; for example, in the physical activity diary, users can track which activity they performed and how much time they spent exercising. Similarly, in the sleep diary, users can track their hours of sleep and factors that might have influenced their sleep quality. In the diary, graphs are made available so that users can keep track of their progress. A notification on their phone prompts users to complete their diaries. The diary is also visible on the dashboard when users log-on to the Minddistrict platform.

In the second and third week of every month, the user evaluates the previous week and is encouraged to identify factors that hindered or facilitated his/her progress.

Goal Evaluation
After 4 weeks, the user revisits the previously created GAS. Using this scale, the user indicates whether and to what extent the goal was reached. This evaluation is followed by personalized feedback. If a goal has been met, the user is asked to set a new goal. If the user has failed to reach a goal, he/she receives tips on how to set a realistic goal for the next month; they can also choose the same goal. Figure 3 shows the cycle of Mijn Fitte Brein that repeats each month.
Behavior Change Techniques
Mijn Fitte Brein facilitates behavior change and stimulates adherence. Behavior change is enabled by multiple methods integrated in the platform. The use of GAS in the goal-setting module helps patients to set realistic and attainable goals [31]. An overview of all integrated behavior change techniques according to the Behavior Change Technique (BCT) taxonomy [32] can be found in Multimedia Appendix 4.

Technical Support
The intervention is completely web based and self-applied, meaning that no involvement of health care professionals is needed for participants to use the program. For questions regarding the use of Mijn Fitte Brein, patients can contact the research team. For technical help users can contact the Minddistrict helpdesk.

Data Logging
The Minddistrict platform automatically stores data such as the number of logins, when a module is started and finished, when a diary is completed, and the total time spent on the platform.

Outcome Measures
Feasibility Study
Primary Outcome Measure: Identification of Areas of Improvement
The patients in the feasibility study will be interviewed after the intervention period. Using semistructured interviews, information about the usability and the patients’ experience of Mijn Fitte Brein will be collected. The interview will cover the following themes: the intervention, lifestyle domains, information on cognition and lifestyle, barriers and facilitators for using the intervention, and whether Mijn Fitte Brein facilitated behavior change. In addition, patients will be asked how they think the intervention could be further improved for patients with cancer. For the interview guideline, see Multimedia Appendix 5.

This feedback will be used to adjust Mijn Fitte Brein for the RCT. The interview will be recorded and summarized. Data will be organized by 2 researchers (AD and MK) to identify areas of improvement. Summaries will be used to gain insight into the user experience of the patients. When patients report difficulty with the use of the intervention, changes will be made to the platform accordingly. When an area of improvement has been suggested more than once and addition to the online intervention is feasible, the platform will be adjusted accordingly before the start of the RCT.

Findings in the log data will also be used to identify changes in adherence and usage of Mijn Fitte Brein. Participants are also asked to clarify reasons for their usage and adherence. These explanations will be used to improve the platform and thereby optimize adherence in the future.

Other Outcome Measures
Usability
The usability of Mijn Fitte Brein will be reported using the Post-Study System Usability Scale (PSSUQ) [33]. The PSSUQ questionnaire consists of 16 items on usability (scores ranging from 1 to 7). The total score ranges from 16 to 112, with higher scores indicating better usability ratings.

Log Data
Log data are used to study the use of Mijn Fitte Brein and adherence.

The use of the platform will be described by the number of logins and the time spent on Mijn Fitte Brein. Furthermore, the number of completed modules, goals, and diaries will be counted and used to further specify the use of Mijn Fitte Brein.

Adherence will be described by reporting the average number of completed sessions, and the number of completed active (diaries/exercised) versus passive (information) sessions. Finally the number of people that never logged in will be reported.

Randomized Clinical Trial
Primary Outcome Measure: Subjective Cognitive Functioning
Subjective cognitive functioning is the primary outcome measure of the RCT. The Dutch version of the FACT-Cog questionnaire [34] will be used to indicate self-perceived cognitive impairment. The questionnaire has high internal consistency and satisfactory test–retest reliability [35]. The FACT-Cog consists of 37 items related to cognitive problems due to treatment in patients with cancer. The items are divided into 4 categories: perceived cognitive impairments (20 questions), comments from others (4 questions), perceived cognitive abilities (9 questions), and impact on HRQoL (4 questions). All items are rated on a 5-point Likert scale ranging from 0 (“Never”
or “Not at all”) to 4 (“Several times a day” or “Very much”). The total score for the FACT-Cog ranges from 0 to 148, with a higher score indicating better perceived cognitive functioning. The threshold for MCID was set at 9.6 points [27].

**Other Outcome Measures**

**Objective Cognitive Functioning**

A neuropsychological test battery is used to measure functioning of multiple cognitive domains, taking approximately 1.5 hours to complete. This test battery is based on the core battery–recommended tests by the International Cognition and Cancer Task Force [36]. These tests are specifically selected for their ability to gauge a broad range of neurocognitive functions in patients with cancer. Additional tests based on previous research and our clinical experience were added to this core battery [37]. The battery consists of widely used standardized psychometric instruments that have been shown to be sensitive to the negative effects on cognition of cancer treatment in other clinical trials. For an overview of the tests, see Table 2. The neuropsychological test data will be scored according to standard scoring algorithms. The included tests have published normative data that take into account age and, where appropriate, education, gender, and handedness. The test results for the patients can also be expressed in a standard form denoting their relation to control data with the use of a z-transformation: $z = (x - m)/SD$, where $x$ is the observed score for the patient; $m$ is the average score for healthy controls; and SD is the standard deviation for healthy controls. This provides measures of functioning which are comparable across tests.

Table 2. Overview of neuropsychological test battery.

<table>
<thead>
<tr>
<th>Test</th>
<th>Cognitive domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hopkins Verbal Learning Test–Revised [38]</td>
<td>Verbal memory</td>
</tr>
<tr>
<td>Trail Making Test A + B [39]</td>
<td>Flexibility</td>
</tr>
<tr>
<td>WAIS-III Digit Symbol Substitution Test [40]</td>
<td>Information processing/attention</td>
</tr>
<tr>
<td>Eriksen Flanker Task [41]</td>
<td>Response inhibition</td>
</tr>
<tr>
<td>Rey Complex Figure [42]</td>
<td>Visuospatial construction and recall</td>
</tr>
<tr>
<td>WAIS-III Digit Span Backwards [40]</td>
<td>Working memory</td>
</tr>
<tr>
<td>Stroop Color Word Test [43]</td>
<td>Attention, response inhibition</td>
</tr>
<tr>
<td>Tower of London [44]</td>
<td>Planning/executive functioning</td>
</tr>
<tr>
<td>Controlled Oral Word Association Test [45]</td>
<td>Verbal fluency</td>
</tr>
<tr>
<td>Behavioral Assessment of the Dysexecutive Syndrome (BADS) 6 Elements Test [46]</td>
<td>Planning/executive functioning</td>
</tr>
</tbody>
</table>

**Health-Related Quality of Life**

The EORTC Core Quality of Life Questionnaire (EORTC QLQ-C30) will be used to establish the patients’ HRQoL [47]. This questionnaire is widely used in clinical trials to evaluate HRQoL in patients with cancer. The questionnaire consists of 30 items, which are distributed over 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea and vomiting), a global HRQoL scale, and additional items to address other common symptoms. The patient is asked to indicate which number on a 4-point Likert scale applies best to them. The total global health status scores range from 0 to 100, with a high score indicating high HRQoL.

**Psychosocial Factors**

Self-efficacy will be measured using the Dutch General Self-Efficacy Scale [48]. This scale consists of 10 items that measure the beliefs that one’s actions are responsible for successful outcomes. Items are scored on a 4-point scale, and the total score ranges from 10 to 40, with a higher score indicating more self-efficacy.

To measure symptoms of fatigue, the Checklist Individual Strength (CIS) will be used [49]. The CIS is widely used for studying fatigue in a working population. This questionnaire consists of 20 items divided over 4 scales: subjective experience of fatigue, concentration, motivation, and physical activity. Items are scored on a 7-point Likert scale. The total score ranges from 20 to 140, with a higher score indicating more fatigue-related problems.

Anxiety and depression state will be measured using the Hospital Anxiety and Depression Scale (HADS) [50]. The scale is widely used to screen for emotional distress in both clinical and research settings. The HADS consists of 2 subscales (each with 7 items). Each item is rated on a 4-point scale from 0 to 3, yielding a maximum score of 21 for each subscale. A higher score indicates more symptoms of anxiety or depression.

One item of the work ability index [51], “current work ability compared with the lifetime best,” will be used to assess total ability to work. This item is suitable to be used as a proxy for the full work ability index [52]. The item has a score between 0 (completely unable to work) and 10 (work ability at its best). To measure cognitive functioning specific for the work context, additional tests based on previous research and our clinical experience were added to this core battery [37]. The battery consists of widely used standardized psychometric instruments that have been shown to be sensitive to the negative effects on cognition of cancer treatment in other clinical trials. For an overview of the tests, see Table 2. The neuropsychological test data will be scored according to standard scoring algorithms. The included tests have published normative data that take into account age and, where appropriate, education, gender, and handedness. The test results for the patients can also be expressed in a standard form denoting their relation to control data with the use of a z-transformation: $z = (x - m)/SD$, where $x$ is the observed score for the patient; $m$ is the average score for healthy controls; and SD is the standard deviation for healthy controls. This provides measures of functioning which are comparable across tests.

One item of the work ability index [51], “current work ability compared with the lifetime best,” will be used to assess total ability to work. This item is suitable to be used as a proxy for the full work ability index [52]. The item has a score between 0 (completely unable to work) and 10 (work ability at its best). To measure cognitive functioning specific for the work context, the Dutch version of the Cognitive Symptom Checklist–Work [53] will be used. This questionnaire consists of 19 items measuring work-specific cognitive problems in patients with cancer. Items are scored on a 4-point scale ranging from 0 (never) to 4 (always). Total scores range between 0 and 100, with a higher score indicating more cognitive symptoms.
To measure work functioning, 2 scales of the Dutch version of the Work Role Functioning Questionnaire will be used [54]. Both the mental and social demands scale, consisting of 7 items, and the flexibility demands scale, consisting of 5 items, will be included. Subscale scores are scored by taking the sum score of the items, divided by the number of items, and then multiplied with 25 to obtain a score between 0 and 100, with higher scores indicating better work functioning.

**Lifestyle Factors**

The lifestyle questionnaires are used as input for goal setting in the Mijn Fitte Brein intervention. There are separate questionnaires for the following domains: alcohol consumption status, physical activity, exercise, healthy nutritional behavior, smoking status, sleep status, and stress status. The questionnaires are based on the recommendations for physical exercise, nutrition, and alcohol use by the Health Council of the Netherlands [55,56]. For a detailed description of the questionnaires, see Multimedia Appendix 6.

**Data Management and Security**

The Minddistrict platform is hosted by Internaxx. Internaxx is SO27001 and NEN-7510 certified. Minddistrict has received an EU Declaration of Conformity (No. 2017/590-01). Data are collected and stored according to the GDPR. All participants are assigned an identification number that is used for the collection of the raw data and in the data analysis.

Data are stored online on the European server of Castor EDC. Paper and pencil tests and informed consent files will be stored separately in a locked file cabinet at Amsterdam UMC, location VUmc, Amsterdam, The Netherlands. Upon completion, questionnaire data will be directly recorded into an online system (Castor EDC). The data will then be transferred into a statistical database for analysis. Raw data will be stored for 15 years.

An on-site monitor has been assigned to the study to make sure the right procedures are followed and the correct documentation is provided before, during, and after the study.

**Data Analysis**

The data analysis plan has been developed in collaboration with a statistician experienced in statistics for clinical cancer research.

**Feasibility Study**

For adherence and use of the intervention, descriptive statistics (eg, percentages, median, and interquartile range) of log data and the PSSUQ will be generated.

Qualitative data collected during the interviews will be summarized and organized by 2 researchers (AD and MK). Data will be categorized by theme to create an overview of possible areas of improvement.

**Randomized Controlled Trial**

**Descriptive Statistics and Demographics**

Descriptive statistics (means [SD]) will be generated for the range of outcome measures and demographics (age, gender, level of education, cancer type, time since diagnosis, time since treatment completion, treatment type, scores on all questionnaires, and scores on neuropsychological tests).

Demographics of the interventions and control group will be compared, including level of education, fatigue (MFI-20 scores), anxiety and depression (HADS scores), using a chi-square test or independent samples t tests where appropriate.

**Effect of Mijn Fitte Brein on Cognitive Complaints Over Time**

To determine the effects of between-subjects factor intervention, within-subjects factor time points (baseline, 3-month follow-up, and 6-month follow-up), and their interaction on cognitive complaints and secondary study parameters, linear mixed models will be used. In case of a significant intervention × timepoint interaction, effect sizes after 3 and 6 months (or a change from baseline) will be computed with an independent samples t test. The assumption of normality will be checked using histograms. In case of violations, nonparametric procedures will be used. Two-tailed probability tests will be used in all inferential statistical testing. A result will be considered significant at P≤.05. A change equal to or greater than the MCID (9.6 points) will be considered as a meaningful change on the FACT-Cog.

**Moderating Effects**

To study the moderating influence of other factors on the effect of Mijn Fitte Brein, a linear mixed model will be used, including between-subjects factors intervention and moderator, and within-subject factor time, all 2-way interactions, and the 3-way interaction. The following possible moderators will be tested: age, gender, self-efficacy (DGSES), cancer type, type of treatment, work performance (Work Role Functioning Questionnaire), level of education, mood (HADS), and fatigue (CIS). Bonferroni-adjusted α levels will be used to control for type I errors due to multiple comparisons.

Statistical analyses will be conducted using SPSS 22 (IBM).

**Bias Due to Loss to Follow-Up and Missing Data**

In both the feasibility study and the RCT, participants who do not complete the study will be questioned about their reasons for dropping out. The responses will be logged and reported.

In the RCT, the number of participants lost to follow-up and their reasons will be compared between groups. In addition, the characteristics of study completers will be compared with the characteristics of noncompleters. Finally, we will analyze missing data to determine whether these are missing at random or not. In the data analysis a linear mixed model will be used, which is able to compensate for missing data.

**Results**

The feasibility study started in February 2020. Because of government regulations for COVID-19, patient recruitment has been deferred for the foreseeable future. As of July 2020, no patients have been enrolled. The findings of the feasibility study will be used to optimize the Mijn Fitte Brein intervention. Enrollment for the RCT will continue when possible. The feasibility study will take 6 months (including making adjustments to the intervention), and the RCT will take 2 years. The final results are expected in 2024. The results of the feasibility study and the RCT will be published in peer-reviewed journals.
Apart from sharing the outcomes of this study with the scientific community, we will approach health insurance companies to make Mijn Fitte Brein more broadly available when it is found to be effective in improving cognition in patients with cancer. The results of the study will also be shared via media such as (local) newspapers and social media.

**Discussion**

This protocol presents a feasibility study and an RCT to investigate whether a multidomain web-based lifestyle intervention, Mijn Fitte Brein, is feasible and effective in a population of patients with cancer returning to work. Cancer-related cognitive impairment is reported in up to 75% of patients with cancer and has a high impact on HRQoL. Lifestyle factors are associated with cognitive functioning in both healthy and cancer populations. The modifiability of lifestyle makes it a suitable target for interventions aimed at improving cognitive functioning in patients with cancer.

By providing patients with Mijn Fitte Brein during the time they are returning to work, the negative impact of cognitive impairment on work performance and experience might be decreased. Furthermore, improvement in lifestyle might benefit (psychological) health and improve HRQoL. Another strength of the study is the 2-phase study design consisting of a feasibility study and an RCT. The findings of the feasibility study will be used to further improve and adjust the intervention for the RCT, resulting in an intervention tailored to the wishes of the patient population. The use of an RCT design makes it possible to control for naturally occurring improvement of cognitive functioning over time and learning effects on the neuropsychological assessment. The assessments at baseline, 3 months, and 6 months make it possible to follow the trajectory of changes in both lifestyle and cognitive functioning.

The study has some possible limitations. First, because of the nature of the study, it is not possible to blind patients to the condition they are assigned to. This might influence test performances. To limit effects of observer bias, test personnel will be blinded to the group the participant is assigned to. Second, there is a chance patients assigned to the control group will have a higher drop-out rate due to not receiving immediate access to the intervention. To prevent this, the control group will be given access to the intervention after the study period. Finally, because the patients are selected by their treating occupational physician, there is a risk of selection bias. However, the occupational physicians are instructed by the researchers to minimize this bias.

This study is the first to investigate the use of a multidomain web-based lifestyle intervention in a cancer population with the goal to improve cognitive functioning. If Mijn Fitte Brein is found to be feasible for our study population and effective in decreasing cognitive complaints in patients with cancer returning to work, it might be implemented in clinical practice. Being both affordable and accessible, Mijn Fitte Brein could contribute to better cognitive functioning and thereby increase the return-to-work rate and quality of life in patients with cancer.

**Acknowledgments**

This study was funded by the AM Tydeman Fonds. We thank Dr B. Lissenberg-Witte for her support in developing the data analysis plan.

**Authors’ Contributions**

AD, MK, and JR were responsible for the design of the study. All authors performed manuscript writing, reviewing, and approval final version of the manuscript.

**Conflicts of Interest**

None declared.

---

**Multimedia Appendix 1**

CONSORT 2010 checklist of information to include when reporting a randomized trial.

[DOCX File, 30 KB - resprot_v10i4e22670_app1.docx]

**Multimedia Appendix 2**

Monthly schedule of MFB.

[DOCX File, 59 KB - resprot_v10i4e22670_app2.docx]

**Multimedia Appendix 3**

Example of the goal setting module.

[DOCX File, 224 KB - resprot_v10i4e22670_app3.docx]

**Multimedia Appendix 4**

Overview of the behavior change techniques used, according to the BCT taxonomy.
Multimedia Appendix 5
Topic list for the semi-structured interviews for the feasibility study.

Multimedia Appendix 6
Description of the lifestyle questionnaires in MFB.

References


**Abbreviations**

- **BAM**: Brain Aging Monitor
- **GAS**: Goal Attainment Scale
- **GDPR**: General Data Protection Regulation
- **HRQoL**: health-related quality of life
- **MCID**: minimal clinically important difference
- **PSSUQ**: Post-Study System Usability Questionnaire
- **RCT**: randomized controlled trial
User-Centered Counseling and Male Involvement in Contraceptive Decision Making: Protocol for a Randomized Controlled Trial

Mahesh Karra¹, BA, MSc, SD; Kexin Zhang², BA

¹Frederick S Pardee School of Global Studies, Boston University, Boston, MA, United States
²Department of Economics, Boston University, Boston, MA, United States

Corresponding Author:
Mahesh Karra, BA, MSc, SD
Frederick S Pardee School of Global Studies
Boston University
152 Bay State Road
Room G04C
Boston, MA, 02215
United States
Phone: 1 617 358 0197
Email: mvkarra@bu.edu

Abstract

Background: To achieve informed choice within the framework of reproductive autonomy, family planning programs have begun to adopt user-centered approaches to service provision, which highlight the individual client as the focal point of interaction and key decision maker. However, little is known about how user-centered approaches to family planning, particularly family planning counseling, shape contraceptive preferences and choices.

Objective: We conducted a multiarmed randomized controlled trial to identify the causal impact of user-centered approaches to family planning counseling on women’s contraceptive decision making in urban Malawi. This study aims to determine how a tailored, preference-driven approach to family planning counseling and the involvement of male partners during the counseling process may contribute to shaping women’s contraceptive preferences and choices.

Methods: Married women aged 18-35 years were recruited and randomly assigned to 1 of the 3 intervention arms or a control arm characterized by the following two interventions: an intervention arm in which women were encouraged to invite their husbands to family planning counseling (husband invitation arm) and an intervention arm in which women received targeted, tailored counseling on up to five contraceptive methods (as opposed to up to 13 contraceptive methods) that reflected women’s stated preferences for contraceptive methods. Women were randomized into a control arm, T0 (no husband invitation, standard counseling); T1 (husband invitation, standard counseling); T2 (no husband invitation, targeted counseling); and T3 (husband invitation, targeted counseling). Following counseling, all women received a package of family planning services, which included free transportation to a local family planning clinic and financial reimbursement for family planning services. Follow-up surveys were conducted with women 1 month after counseling.

Results: A total of 785 women completed the baseline survey, and 782 eligible respondents were randomized to 1 of the 3 intervention groups or the control group (T1, n=223; T2, n=225; T3, n=228; T0, n=108). Furthermore, 98.1% (767/782) of women were contacted for follow-up. Among the 767 women who were contacted, 95.3% (731/767) completed the follow-up survey. The analysis of the primary outcomes is ongoing and is expected to be completed by the end of 2021.

Conclusions: The results from this trial will fill knowledge gaps on the effectiveness of tailored family planning counseling and male involvement in family planning on women’s stated and realized contraceptive preferences. More generally, the study will provide evidence on how user-centered counseling may affect women’s willingness to use and continue contraception to realize their contraceptive preferences.


International Registered Report Identifier (IRRID): DERR1-10.2196/24884

https://www.researchprotocols.org/2021/4/e24884

JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e24884 | p.68

(page number not for citation purposes)
user-centered counseling; male involvement; contraceptive use; family planning; randomized controlled trial; Malawi; Sub-Saharan Africa

Introduction

Choice Architecture in Family Planning

Unlike other domains in women’s health, delivering high-quality family planning services is reflected by the achievement of good reproductive health outcomes and the objective of helping women maximize a complex set of preferences around future fertility outcomes. In family planning, the role of the client as the key actor in her receipt of care is distinct from most other contexts in health-related decision making, wherein providers often play a leading role in determining which course of treatment is best for a patient. Indeed, family planning programs typically consider women to have the right to full information about contraceptive method options. For this reason, family planning programs dedicate significant resources to providing complete and accurate information so that women are informed of the full range of contraceptive methods. Clients typically do not receive modern contraceptive methods without receiving a comprehensive consultation session with a family planning counselor (typically a health care professional or social worker), during which time they are informed about the range of available methods. Due to the high value placed on fully informed choice, counselors may discuss as many as 13 different methods and describe as many as 10 method attributes (eg, effectiveness at preventing pregnancy, the ease of use, the risk of side effects, and the duration of effectiveness, among others) for each method with a client. This information-intensive approach to counseling, which compels a client to interpret a large volume of information across several dimensions (attributes and methods), may reduce the salience of counseling while simultaneously increasing the potential for choice overload.

Several studies have examined how family planning counseling affects women’s contraceptive uptake [1]. However, little is known about how the choice architecture for family planning—the structures and processes by which contraceptive methods and information are presented to women—actually shapes women’s preferences. Studies have shown that a woman’s fertility intentions, which affect her contraceptive preferences, are likely to be unstable over her reproductive lifetime [2-4] and are sensitive to relatively small changes in her environment. A number of studies have examined the effect of choice architecture on people’s decision-making processes [5-7] and demonstrated that increased cognitive load leads to more risk-averse behaviors, higher levels of impatience, and a higher likelihood of deferring decision making. In addition, studies have revealed that choice architecture—based approaches can be used to nudge an individual to make better choices without forcing certain outcomes [8-12]. For example, when faced with a small number of well-understood alternatives, a person tends to examine all the attributes of all the alternatives and then make trade-offs when necessary. However, when the choice set becomes large, alternative strategies (eg, structuring complex choices into a certain order) are employed [13]. These studies suggest that the complexity of options may adversely affect an individual’s choice. In the context of family planning, Delavande [14] showed that the extent to which a contraceptive method is effective, the extent to which such a method protects against sexually transmitted infections (STIs), and the extent to which the method is approved (or disapproved) by a woman’s partner are the three most important method-related attributes that drive a woman’s choice of contraception. As a woman makes her contraceptive choice based on the attributes of a certain method, information about methods that are tailored to a woman’s most valued attributes may, in turn, allow her to reinforce and better realize her preferences for a contraceptive method.

Male Involvement in Family Planning Counseling

In addition to the role of choice architecture, the involvement of men in family planning and reproductive health decision making, specifically partner involvement in contraceptive counseling, also plays a crucial role in shaping women’s preferences. In most family planning counseling programs however, male partners are often not required to be counseled. Often, male partners do not participate in the counseling process unless they are actively brought by women themselves. A wide range of studies have shown that including men in family planning counseling might increase the use of contraceptive methods through two channels [15-20]. First, counseling may provide men with more detailed information on the benefits of family planning and contraception [15,16]. Second, couples counseling provides a platform for increased spousal communication and offers couples the opportunity to discuss fertility and method preferences more openly. This result is attested in a series of cross-sectional studies that found a positive link between spousal communication and contraceptive use [21-25]. A small number of impact evaluations of male involvement in counseling have been conducted in a handful of countries, including China [26], Ethiopia [27], Malawi [15], Zambia [28], Tanzania [21], and Jordan [29]. To date however, experimental evidence on spousal concordance and the role of men in family planning decision making remains limited and mixed, particularly in low- and middle-income settings. Moreover, existing family planning programs cannot compel women to invite their husbands or male partners to be counseled; therefore, couples who were enrolled for counseling were, by construction, selected because they were willing to be counseled together. These couples and, in particular, the male partners who were selected and participated in counseling are not likely to be representative of typical couples and men. To test the impact of male involvement more effectively, it would be necessary to examine the extent to which giving women the choice to invite or involve their husbands impacts male involvement and subsequent outcomes.
User-Centered Approaches to Family Planning

To achieve full, free, and informed choices within the framework of reproductive autonomy and rights, programs have increasingly begun to adopt user-centered approaches to family planning counseling and service provision. These approaches have stressed the role of the individual client as the focal point of interaction and the key decision maker. Recently, user-centered counseling approaches to family planning, such as the Population Council’s Balanced Counseling Strategy (BCS) [30,31], IDEO’s human-centered design methods [32,33], and others, have been developed to allow counselors to elicit women’s contraceptive preferences along various dimensions more effectively. In developed countries, user-centered counseling strategies on family planning have also been implemented. A decision support tool named My Birth Control—which has educational and interactive modules and produces a provider printout with the patient’s preferences—has been proven to improve patient centeredness of contraceptive counseling [34]. These strategies were designed to allow for the identification of a more tailored range of contraceptive methods that better match women’s preferences. Such counseling approaches, although promising, have some limitations. For instance, many of them are found to be time-consuming and difficult to scale up to larger client bases. A recent report on the BCS shows that the BCS toolkit includes (1) an algorithm that summarizes 11 steps needed to implement the strategy; (2) counseling cards with basic information about 15 family planning methods, plus a card with the checklist to ensure that a woman is not pregnant; and (3) brochures on each of the methods for clients. In addition, there is little evidence on the effectiveness of these approaches in meeting women’s fertility desires [35].

Family Planning Counseling in Malawi

Counseling with a service provider is often the first step for women to learn about, choose, and receive family planning services in Malawi. As stated in the National Reproductive Health Service Delivery Guidelines 2014-2019, counseling is intended to be an interactive process in which the provider seeks to identify the client’s needs, elicit the client’s concerns, and offer relevant information and guidance to enable the client to make an informed decision about methods. During the counseling session, women often receive a range of information, such as the cost of procuring contraceptives (although many methods are provided for free in public facilities), side effects or contraindications associated with methods, and method effectiveness [36]. In public health facilities in Malawi, women typically receive a group counseling session with a nurse or family planning counselor, followed by a short (an estimated 3- to 5-min) individual counseling session, at which time they may choose to receive a contraceptive method. According to the guidelines set by the Ministry of Health (MOH) and the Malawi Reproductive Health Directorate (RHD), a family planning counseling session is typically administered to women with a family planning flip chart, which describes 13 contraceptive methods that are organized in order of method effectiveness to prevent pregnancy, starting with female and male sterilization, and concluding with traditional methods of contraception (Multimedia Appendix 1). Although the flip chart is comprehensive in the information provided to each woman about each method, this counseling procedure does not prioritize women’s preferences for either method attributes or methods themselves. Moreover, counseling in this manner may likely anchor women’s preferences to methods based on a default top attribute (method effectiveness), which may not be the most preferred. Given the limited time for individual counseling, there is also little opportunity for women to receive clarification or follow-up that they may seek, and service providers may not be able to fully elicit a woman’s family planning and fertility preferences before providing them with the most suitable services. Finally, most counseling sessions, particularly those involving group counseling, are exclusively targeted to women clients, with few opportunities for men to participate in the service provision and decision-making processes for family planning.

Study Objectives

In this study, we identified the causal impact of user-centered approaches to family planning counseling on women’s contraceptive preferences and decision making by means of a randomized controlled trial. In particular, we investigated two channels that have been hypothesized to play a role in contraceptive decision making: (1) a preference-based, targeted, user-centered approach to family planning counseling and (2) male involvement in family planning counseling. To this end, we tested the following 2 hypotheses:

1. A preference-driven, tailored, user-centered approach to counseling would allow women to express and realize their contraceptive preferences more effectively.
2. The involvement of male partners in family planning counseling may allow women to express their contraceptive preferences and, in turn, translate their preferences into behavior more effectively.

The study population comprised married women aged 18-35 years living in Lilongwe, Malawi. As part of the trial, each woman in the study was randomly assigned to 1 of the 3 treatment arms or a control arm. A woman assigned to one of the intervention arms received one tailored, user-centered counseling session, followed by a package of family planning services that was designed to reduce the key barriers to accessing family planning in urban Malawi. The primary objective of this study is to evaluate the impact of user-centered counseling on a woman’s contraceptive use, intention to change stated ideal methods, switching of the current method, and realization of the stated ideal method. Primary outcomes include changes to women’s contraceptive method preferences (both stated and realized) over time and changes to women’s contraceptive concordance (whether a woman’s stated choice of contraceptive method, her ideal method, is the method that she is, in fact, using). Secondary outcomes include changes to male partners’ fertility and family planning preferences, women’s sexual and marital well-being, and changes in women’s decision-making in the household.

This study seeks to fill the current knowledge gaps in the evidence on the causal impact of user-centered counseling on women’s contraceptive decision making and behavior. Specifically, the study documents how women’s preferences
for family planning may change over time and investigates the extent to which women’s stated contraceptive preferences are realized and translated into behavior. A downstream objective of this study is to investigate the extent to which improved counseling and outreach may contribute to women's empowerment and autonomy, more generally.

**Methods**

**Study Approval**

Human subject approvals were received from the Boston University (BU) Institutional Review Board (IRB; IRB Protocol Number: IRB5162E), the Malawi National Health Sciences Research Committee (NHSRC; NHSRC Approval Number: 2350), the Lilongwe District Council, the Malawi Police Service, and the Malawi MOH to conduct the study. A Memorandum of Understanding was established with the Good Health Kauma Clinic.

**Study Setting**

Our study was conducted in urban areas of Lilongwe, the capital of Malawi. Estimates from the 2015-2016 Malawi Demographics and Health Survey (MDHS) demonstrate that the contraceptive prevalence rate in Malawi was 45.2% among all women of reproductive age (aged 15-49 years) and 59.2% among married women of reproductive age [37]. These estimated contraceptive prevalence rates exhibit a significant increase from the 2010 MDHS; nevertheless, the unmet need for family planning has remained high, with an estimated 18.7% of women in Malawi reporting to have an unmet need for spacing or limiting births. Injectable contraceptives were the most popular method in Malawi in 2010 and were used by 22.5% of women, followed by intrauterine devices and female sterilization at 9% and 8.3%, respectively [38]. The distribution of contraceptive users by method, or contraceptive method mix, has not changed significantly over time among married women in Malawi, as injectable contraceptives, intrauterine devices, and female sterilization remain the most popular methods used by 30%, 11.5%, and 10.9%, respectively. In Malawi, nearly 8 in 10 (79%) modern contraceptive users aged 19-45 years procure their method from public sector providers (government hospitals or clinics), whereas 8% of users procure their method from nongovernmental providers such as Banja la Mtsogolo (BLM), 6% of users procure their method from the private medical sector, and 4% of users procure their method from the Christian Health Association of Malawi and other faith-based service providers [37].

Although the contraceptive prevalence rate in Malawi has continued to rise over the past decade, the contraception discontinuation rate in Malawi has also remained high, with more than 37% of women reporting that they discontinued their family planning method within the last year, among which half of them discontinued because of nonfertility-related reasons (eg, method-related reasons, such as side effects). This high rate of contraceptive discontinuation suggests that barriers may exist in a woman’s decision-making process that prevent her from choosing the right method that caters to her specific preferences. Although an increasing number of family planning programs have been successful in increasing the contraceptive uptake, it is crucial to be reminded that a woman’s family planning preferences are not realized simply from an increased use of contraceptive methods, which has been reiterated in numerous settings by reproductive rights researchers, policy makers, and advocates alike [39]. Family planning programs have mainly focused on the extensive margin of family planning utilization (ie, increasing contraceptive uptake); however, few existing studies have focused on the intensive margin to determine whether the increase in utilization implies that women’s preferences, specifically her choice of contraceptive method, are, in fact, being met.

**Study Sample and Inclusion Criteria**

This study is a multiarmed randomized controlled trial that was conducted with a sample of women of reproductive age from Lilongwe, Malawi. The study comprised a baseline survey, randomization into 1 of 3 treatment arms or 1 control arm, and the implementation of a 2-month intervention. One follow-up survey was conducted after the intervention and was administered to women either at the clinic or on the phone. If participants could not be reached in either way, then a home visit was paid to the households. Figure 1 outlines the general framework of the entire field experiment.
For this study, we recruited 782 women who, at the time of the baseline survey, (1) were married, (2) were aged between 18 and 35 years, (3) lived in the city of Lilongwe (permanent residents), (4) were not pregnant and did not give birth in the 6 months before the initial screening, (5) had neither been sterilized nor had a hysterectomy, (6) had given birth to at least one child (one live birth) in their lifetime, and (7) lived with their husbands at the time of the screening. These criteria were designed to identify women who were most likely to seek and use a contraceptive method if given the opportunity and means. Women who successfully met these seven inclusion criteria and who consented to participate in the main part of our study were recruited. In addition, no two eligible women were enrolled from the same household. If multiple women from the same household were potentially eligible to be recruited, then we chose the youngest eligible woman from the household to participate. Given that randomization was to be administered at the individual woman level, it was necessary to select only one eligible woman from a household to minimize any possible contamination across the intervention and control arms. We also ensured that eligible women who were selected for the study were sufficiently distant (at least five households apart) from each other, which also served to reduce any spillover effects between treated and control women who lived in the same neighborhood.

In addition, we interviewed the husbands of women who were invited to the counseling session by their wives. The men who were selected to be interviewed were husbands who (1) were assigned to 1 of the 2 intervention arms that encouraged women to invite them and (2) consented to participate in the counseling session with their wives.

### Recruitment and Study Timeline

Our study recruitment design closely follows the recruitment protocols of Karra and Canning [40]. Using the most recent Demographics and Health Survey (DHS) and census maps of Lilongwe’s enumeration areas and listings of households and neighborhoods, we employed a 2-stage sample selection procedure that is based on the sampling strategy used by the DHS. In the first stage, we randomly selected defined census areas in Lilongwe to be screened until we have selected enough enumeration areas to contain at least 5000 households in total. In the second stage, our enumerators proceeded door-to-door to screen households in each selected enumeration area for potentially eligible women. Enumerators screened each household that they approached to determine if any woman who was living in that household (1) met our inclusion criteria listed in the section above and (2) consented to participate in our study. To make this determination, enumerators used a recruitment script to verify eligibility and presented the eligible woman with a consent form to participate in the study. Multimedia Appendix 2 presents the recruitment script and the consent form. Written (or verbal) informed consent was obtained from all participating women before proceeding to administering the survey. On the basis of our knowledge of participation refusal rates and the estimated proportion of eligible women in Lilongwe, we estimated the need to screen at least 940 households across our randomly selected enumeration areas to attain a desired sample size of at least 700 women. We also needed to choose enough enumeration areas to have at least 5000 households in total (940×5=4700 households among the women who made up our sample and who were at least five households apart, plus an additional 300 households that were screened, but where women either did not meet the eligibility criteria or refused to participate in the counseling session with their wives.}

![Figure 1. Experimental framework and flowchart.](image-url)
participate). Recruitment from the selected enumeration areas ceased after at least 700 women were found who met the eligibility conditions, consented to participate in the study, and were administered the baseline survey.

This study spanned a total of 14 months from January 2019 to February 2020. Table 1 presents the study timeline. Before recruiting and conducting the study with the main study sample, a small sample of women was recruited to pilot new survey instruments and intervention activities over the study period. All research activities, including recruitment, consent, study instrument administration, and intervention administration, were administered to the pilot sample using the same study protocols as those used for the main sample. For this reason, the final analytic data sets for the study comprised data from both the pilot and main samples.

Table 1. Study timeline.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2019</td>
</tr>
<tr>
<td></td>
<td>Months 1-5</td>
</tr>
<tr>
<td>Preparation of survey instruments, consent forms, recruitment materials, and Boston University Institutional Review Board protocols</td>
<td>✓</td>
</tr>
<tr>
<td>Hiring of study staff and interventionists, obtaining local institutional review board approval in Malawi, and identifying sampling strategy</td>
<td>—</td>
</tr>
<tr>
<td>Training of local staff, lay interventionists, and enumerators</td>
<td>—</td>
</tr>
<tr>
<td>Baseline survey administration, randomization, and intervention administration</td>
<td>—</td>
</tr>
<tr>
<td>Intervention period—counseling</td>
<td>—</td>
</tr>
<tr>
<td>Follow-up surveys at the clinic through phone calls and through visits</td>
<td>—</td>
</tr>
<tr>
<td>Final reports, publications, and dissemination</td>
<td>—</td>
</tr>
</tbody>
</table>

a: The specified study activity was conducted in the time period.
b: Not available. The specified study activity was not conducted in the time period.

Informed Consent and Participant Privacy

Our consent and privacy processes closely follow the consent protocols established by Karra and Canning [40]. The process of obtaining consent was consistent for all potential participants: women and eligible husbands. Written consent to participate was obtained from all the participants before administering the surveys. Once participants agreed to join the study, a copy of the consent form script was read to them, and they were given opportunities to ask questions and express concerns. Surveyors checked for comprehension throughout the consent process, which took 5-10 minutes on average. Multimedia Appendix 1 presents the consent documents and details of the consent process. After completing the consent script, potential participants were asked if they would like to participate. If the participant wished to take further time to reflect, then the surveyor and the participant would determine the time and method of reconnecting. If the potential participant agreed to participate in the study, then consent was obtained and documented by obtaining the signatures of both the participant and the study staff member who provided consent. Upon providing consent, the surveyor would then conduct the survey.

For follow-up surveys, written consent to participate was obtained from women who were followed up either at home or at the clinic, and verbal consent to participate was obtained over the phone from participants who were originally recruited at baseline but who did not visit the clinic over the intervention period. Respondents who had previously indicated that they were no longer interested in participating in the study were not contacted. All survey responses were collected on Android (Google LLC)-based tablets, and data were securely transferred to the CommCare-supported server daily.

Throughout the consent process, each participant was informed that even if she decides to participate and sign the consent form, she could decide at any time to end her participation. If the prospective study participant was not literate, then a witness who did not work for the study signed the consent form. In the absence of witnesses, the participant could confirm their consent by placing a thumbprint on their consent form, and a photo of the thumbprint was taken as a record of consent. Moreover, participants were encouraged to contact the researchers with further questions at any time during the study. Consent was obtained before any surveys were conducted, and all participants were informed that their participation was completely voluntary.

All recruitment scripts, consent forms, and survey instruments were translated by a certified translator from English into Chichewa and were back-translated into English by a second certified translator to ensure accuracy.

When presented with the intervention, women were offered each intervention component (counseling and family planning...
package) by an interventionist from the study team. Women were informed that they could take up and stop any or all components of the intervention at any time. Counselors asked for a woman’s consent to participate before each counseling session.

All surveys and administration of the intervention were conducted in a private room. To maintain a respondent’s privacy during an attempt to reach her by phone, the field enumerator conducting the call left no indication of the reason for the phone attempt (eg, voicemail and text message) on the respondent’s phone should there be no response to a call. To ensure that the respondent’s participation remains private, the field enumerator only continued with the phone call once she received assurance from the woman that she was able to speak on the phone without being overheard or interrupted. Any disruption or interruption during the phone call resulted in postponement or termination of the call.

Randomization

Following the baseline survey, women who consented to participate were randomized into 1 of the following 4 experimental arms:

1. A control group (T0), in which a woman received a private counseling session on the full range of 13 contraceptive methods following a standard counseling process (n=100).
2. An intervention group (T1), in which a woman was encouraged (but not compelled) to invite her husband to a joint counseling session. A woman (and her husband if she chose to invite him) then received a private counseling session on a full range of 13 contraceptive methods (n=200).
3. An intervention group (T2), in which a woman received a private counseling session on 5 targeted contraceptive methods based on her baseline preferences for family planning (n=200).
4. An intervention group (T3), in which a woman was encouraged to invite her husband to a joint counseling session. A woman (and her husband if she chose to invite him) was then counseled on 5 targeted contraceptive methods based on her baseline preferences for family planning (n=200).

Upon the completion of baseline data collection, women were randomized to 1 of the 3 intervention arms or the control arm, such that the intervention assignment was balanced along the following baseline characteristics: the use of family planning, most preferred method attributes, and age. As part of the balancing process, strata by each combination of characteristic values were created, and observations were assigned to their respective strata. Observations within each stratum were then individually randomized to each of the 4 groups by the principal investigator (PI) in a 2:2:2:1 allocation using computer-generated randomization using STATA, version 14 (StataCorp LLC). Following individual randomization, the Innovations for Poverty Action (IPA) Malawi intervention team implemented the intervention in the participants randomized into the intervention arms. Although the implementation team was not involved in the actual randomization process, counselors, who were responsible for administering the correct counseling and follow-up procedures to each respondent, were not blinded to respondent assignments to each of the intervention arms.

The Intervention

Intervention Protocols

All women who consented to participate in the study (T0-T3) were offered a free, private family planning counseling session, which included a risk assessment for contraceptive methods and detailed information on methods switching, side effects associated with each method, and the benefits of contraception and birth spacing. A precounseling survey was administered to each woman, where they were asked to confirm their personal information, including, but not limited to, pregnancy status, current contraceptive use, and preferences for fertility and contraception. During counseling sessions, counselors provided women with information on the range of modern and natural family planning methods based on their respective intervention assignments. Strategies on how to communicate family planning messages with partners and on how to increase partner awareness were conveyed during the sessions. The counseling session lasted no more than 1 hour and was administered in a private room by a counselor who was trained to provide family planning and reproductive health services. To facilitate an effective and open dialog with women during the counseling session, only female counselors were hired. Counselors were hired by IPA Malawi, held a nursing degree, and were trained on the range of counseling protocols over a week. We enlisted the support of the Malawi RHD and several international nongovernmental organizations who work on family planning, including Population Services International (PSI), BLM, the Family Planning Association of Malawi, and FHI360, to help us develop training materials and counseling resources. We also collaborated with the Malawi RHD, BLM, and PSI to assist with counselor training.

Counseling for Control Group T0

Following the introductory risk assessment and discussion on the benefits of contraception and birth spacing, women who were assigned to the control arm were counseled on the full range of 13 available family planning methods. Counselors employed the standard-of-care contraceptive method flip chart that is provided by the MOH and RHD and counseled women on each method following the order of methods in the flip chart.

Counseling for Intervention Group T1

Before receiving the counseling session, women who were assigned to the intervention arm T1 were encouraged by the counselor to invite their husbands or male partners to participate in a joint family planning counseling session. Following the invitation, women and their husbands (if they chose to invite them) jointly received an introductory risk assessment and discussion on the benefits of contraception and birth spacing. Women and their husbands were then jointly counseled on the full range of 13 available family planning methods using the standard counseling flip chart, as described for intervention arm T0.
Counseling for Intervention Group T2

Following the introductory risk assessment and discussion on the benefits of contraception and birth spacing, women who were assigned to the intervention arm T2 were counseled on a targeted number of methods that were chosen based on the respondents’ reported contraceptive preferences at baseline. The objective of this intervention arm was to minimize the choice overload and increase the salience of a woman’s most preferred method attribute (e.g., method effectiveness in preventing pregnancy, duration of use, and likelihood of method-related side effects). At baseline, each woman was asked to assign a relative rank to her top 3 most valued method-specific attributes (e.g., does she prefer that a method has a lower incidence of side effects over a method that is more effective at preventing pregnancy?). On the basis of her ranking of method attributes, the counselor confirmed the attribute that the woman revealed to be most important (e.g., methods with low incidence of side effects) and used a predesigned tailored flip chart (an abbreviated version of the full flip chart) to present a subset of 5 methods that ranked highest along that revealed attribute. Particular emphasis was placed on making the order of presentation salient, in which women were reminded and primed to consider the relative ranking of methods along the stated attribute. Counselors then counseled women on each of the 5 methods following the standard family planning counseling procedure described above. Additional details on each of the predesigned tailored flip charts are presented in Table 2.
<table>
<thead>
<tr>
<th>Flip chart color and methods</th>
<th>Attributes</th>
</tr>
</thead>
</table>
| **Blue**                     | • Effective at preventing pregnancy  
• Duration of effect or lasts long |
| Sterilization                |            |
| IUD                         |            |
| Implants                     |            |
| Injectables                  |            |
| Pill                         |            |
| **Purple**                   |            |
| LAM                         | • No risk of harming health |
| Two-day method               | • No effect on monthly bleeding |
| Rhythm method                | • No unpleasant side effects |
| Standard days method         | • Low cost |
| Condoms                      | • No risk of infertility |
| N/A                          | • Nonhormonal |
| N/A                          | • No need to go to the clinic to procure the method |
| **Pink**                     |            |
| Condoms                      | • Immediate return to fertility |
| Two-day method               |            |
| Rhythm method                |            |
| Standard days method         |            |
| IUD                          |            |
| **Yellow**                   |            |
| Condoms                      | • Protects against HIV or STId |
| **Gray**                     |            |
| IUD                          | • Want to try something new or tired of old method |
| Implants                     | • My doctor recommended it to me |
| Sterilization                | • My husband wanted me to use this method |
| Pills                        | • Other women in my family have used this method |
| Injectables                  | • Friends have used this method |
| N/A                          | • Easily available at clinic |
| **Orange**                   |            |
| Sterilization                |            |
| IUD                          |            |
| Implants                     |            |
| Injectables                  |            |
| **White**                    | • No need to go to the clinic to resupply |

Table 2. Attribute-method flip chart correspondence.
Counseling for Intervention Group T3

Before receiving the counseling session, women who were assigned to the intervention arm T3 were encouraged by the counselor to invite their husbands or male partners to participate in a joint family planning counseling session. Following the invitation, women and their husbands (if they chose to invite them) received an introductory risk assessment and discussion on the benefits of contraception and birth spacing. In following the counseling protocols of the T2 intervention arm above, a woman and her husband were jointly counseled on a targeted range of 5 family planning methods based on the woman’s highest ranked preferred method attribute she identified at baseline. Before counseling, counselors confirmed the woman’s highest ranked attribute and jointly counseled the woman and her husband on the targeted subset of 5 contraceptive methods that most closely aligned with her most preferred attribute using a tailored, condensed flip chart specified in Table 2.
Postcounseling Survey

Survey Instrument and Rollout

Following the counseling session, counselors conducted a brief survey with all women to assess their experiences with the counseling service. Multimedia Appendix 2 presents the postcounseling survey instrument. A particular aim of this follow-up survey was to elicit women’s preferred choice of contraceptive method and the cited reasons for choosing this method immediately after counseling. Interviews with women and husbands who were present took no more than 15 minutes and were conducted privately in accordance with the standard interviewing procedures that were established for the baseline survey.

Postcounseling Package of Services

Following the postcounseling survey, all women (and husbands who participated in the counseling session) were offered the following package of services for a 1-month period:

1. Transportation service: women (and applicable husbands) were offered a free transportation service from their homes to the Good Health Kauma Clinic for a period of 1 month. The transportation service was provided by a driver who was hired and trained by IPA Malawi. Respondents received the phone number of a female field manager and were instructed to let the field manager contact the driver to transport them to the Good Health Kauma Clinic during the clinic’s normal working hours, that is, from 8 AM to 5 PM from Monday to Saturday. The field manager maintained a daily schedule of respondents who requested the driver’s services, and the respondents were instructed to notify the field manager at least one day before they wish to go to the clinic to ensure that the driver would be able to transport them. The field manager also provided 1 day’s advanced notice to the Good Health Kauma Clinic to inform them of the number of clients who could be expected to attend the clinic on the following day. Each woman who visited the clinic was always accompanied by the female field manager; the presence of another woman in the vehicle acted to minimize potential stigma associated with a woman traveling alone in the company of a man and also provided comfort to the participant.

2. Financial reimbursement for family planning services: finally, women and participating husbands were financially reimbursed for any out-of-pocket expenditures that they incurred for receiving family planning care at the Good Health Kauma Clinic for the 1-month service period. Costs that were eligible for reimbursement included the procurement costs of family planning medications and contraceptive methods, family planning consultation fees, lab test fees, treatment costs for any contraceptive-related side effects and contraindications, expenses associated with switching and discontinuation of methods, and exam fees. Each couple was allowed a maximum reimbursement amount of MWK 17,500 (US $25.00), which could be redeemed by the couple over multiple visits at the Good Health Kauma Clinic. The extent to which a given expenditure at the clinic met the criteria for being reimbursed was determined by the clinician. For every family planning service that was eligible for reimbursement, the cost of the service was deducted from her reimbursement allowance.

In addition, women and participating husbands who experienced any side effects because of contraceptive use over the course of the service period were entitled to free treatment. If a woman or her husband experienced a side effect or contraindication, then she or he may seek care at her nearest public clinic, public hospital, or the Good Health Kauma Clinic. The woman was asked to keep receipts of any costs incurred at the health facility so that she could be reimbursed later. The reimbursement covered the cost of side effects treatment for all family planning methods used by the woman, regardless of where the method or treatment was procured. All reimbursements for incurred costs were distributed as closely as possible to the time that the reimbursable cost was incurred, most likely within 2-3 business days by the field manager.

At the end of the counseling session, the family planning counselor presented a terms of service document, which specified the terms and conditions for each of the abovementioned services to the woman. Multimedia Appendix 2 presents the terms of service documents for each of the 4 interventions. Counselors answered any questions from the woman about the package of services and then asked for the woman’s (and her husband’s) understanding of these terms. Each respondent also received a paper copy of the terms of the service document.

Follow-up

Follow-Up Phased Rollout

Following the 1-month service period, the entire study sample of women was resurveyed with an abbreviated version of the baseline survey questionnaire. Resurveying participants in this manner enables us to create a panel of individual women where each woman is observed at 3 time points. The same protocols regarding data security and confidentiality that were enforced at baseline were implemented once again for each follow-up survey round.

Follow-up surveys were administered in 3 phases: (1) a clinic-based survey that was administered to women (and participating husbands) who visited the Good Health Kauma Clinic, (2) a phone follow-up survey that was administered to women who did not visit the Good Health Kauma Clinic, and (3) an in-person home-based follow-up survey that was administered to women who neither visited the Good Health Kauma Clinic nor were available for a phone follow-up survey.

Clinic-Based Survey

The 1-month clinic-based follow-up survey was administered by the local field team and in person to all women (and husbands) who visited the Good Health Kauma Clinic over the 1-month service period. The clinic-based survey, which is an abbreviated version of the main baseline survey instrument, took no more than 20 minutes and was administered by the female field manager who accompanied the woman to the clinic.
Phone Follow-Up Survey
For women who either moved out of Lilongwe or did not visit the Good Health Kauma Clinic, the field team attempted to contact them by phone up to a maximum of 3 times. Women who were reachable by phone were invited to participate in a short phone follow-up survey. To maintain a respondent’s privacy during an attempt to reach her by phone, the field enumerator conducting the call left no indication of the reason for the phone attempt (e.g., voicemail, SMS text message) on the respondent’s phone if there was no response to a call. The phone follow-up survey instrument took no more than 20 minutes and was identical to the clinic-based follow-up survey instrument.

Home-Based Survey
For women who did not visit the Good Health Kauma Clinic and who were also uncontactable by phone, the field team attempted to contact them at their homes up to a maximum of one time. Women who were reachable at their homes were invited to participate in a home-based follow-up survey, which was administered in person by a field enumerator. The home-based follow-up survey instrument took no more than 20 minutes in total and was identical to the clinic-based and phone-based follow-up survey instruments. The survey was administered in a private room at a woman’s home.

All protocols related to the consent process and in-person survey administration followed during the baseline survey were also followed for the clinic, phone, and home surveys. Before administering each survey, the field enumerators clearly explained the purpose of the follow-up to the respondents and asked for their consent to participate. For in-person clinic-based and home-based surveys, written consent from respondents was solicited, whereas verbal consent was obtained from participants who were contacted for the phone survey. Following the receipt of consent, the enumerators would then ensure the confidentiality and privacy of responses by asking the respondent to find a private room or space in their household (for the baseline, counseling, phone follow-up, and home follow-up surveys) or at the clinic (for clinic-based surveys) where their responses could not be overheard. For each woman, survey data were collected up to a total of 3 time points (at baseline, at the counseling stage, and at the follow-up stage).

Participant Compensation
All women who participated in the study received a small token of appreciation (3 bars of soap, a monetary equivalent of MWK 500 [US $0.66]) after completing the baseline survey. No monetary compensation was provided to participants for participating in the study, which served to minimize coercion. Participants were not informed of these tokens until the baseline survey had been administered.

All participants received a mobile phone credit of MWK 100 (US $0.14) by means of a mobile airtime transfer. To avoid coercion into participation, women were not informed of this airtime transfer until after the counseling session had been administered.

Primary Outcomes
Key primary outcome variables include the following:
- Attitude or knowledge of family planning, including knowledge of family planning and perceptions toward contraception (e.g., intentions to use)
- Contraceptive use, including changes in contraceptive use, changes in method mix, and adherence to methods (compliance, switching, and discontinuation)
- Contraceptive preferences, including the ideal contraceptive method and the most valued contraceptive method attribute
- Pregnancy and fertility outcomes, including pregnancy status, parity, delivery in a facility, months since last birth, wantedness of last birth, and intentions to become pregnant in the future
- Use of local health facilities, including the frequency of visits to local clinics and frequency of receipt of family planning counseling services
- Husband’s preferences, including the approval of family planning and fertility preferences
- Women’s autonomy, empowerment, and decision making

Secondary Outcomes
A range of secondary outcomes were collected in each survey wave: (1) women’s education, employment, work, and income; (2) household assets; (3) postpartum treatment from last birth; (4) sexual and marital well-being; (5) husband’s employment, work, education, and earnings; (6) household bargaining and women’s autonomy and empowerment; and (7) trust in local health facilities. Outcomes were collected according to the schedule outlined in Table 3. All survey instruments (baseline, counseling, and follow-up) and intervention monitoring tools to track participants and intervention uptake over the course of the study are available upon request.
Survey Instruments and Monitoring Tools

A paper version of the baseline survey instrument, the counseling questionnaire, the clinic-based survey instrument, the home-based survey instrument, and the phone survey instrument are provided in Multimedia Appendix 2. Baseline surveys were administered in a private room in the woman’s place of residence and lasted approximately 2 hours. Short breaks of 5 to 10 minutes were given to the respondents at the end of each section, and additional breaks were taken at the request of the participant and at other scheduled times (eg, mealtime and picking up children from school) as needed. Surveys were conducted in Chichewa, the local language, and strictly follow the format of the questionnaires, which were electronically programmed into Android tablets. To successfully track our participants over time, we collected identifiable data, including names and contact information, names and background characteristics of other household members, household addresses, contact phone numbers and emails, and GPS coordinates of the household. For the purposes of minimizing loss to follow-up, we took photographs of the household and of the survey respondents, and we asked respondents to provide the names and contact information of 2 contacts who did not live in the household and who could be contacted if the respondents could not be directly located in the follow-up period. To protect participant privacy and ensure confidentiality, all identifiable data are appropriately stored and secured in accordance with the data security measures described in the Data Safety and Monitoring section below.

Analysis Plan

The analysis of quantitative study data will be conducted using STATA and R (R Foundation), where appropriate. Descriptive analysis will be performed for all variables, and unadjusted comparisons between experimental arms will be conducted. Descriptive statistics, including frequencies, means, and SDs, will also be performed. Chi-square tests and one-tailed t tests will be used to examine associations in the data. A probability value of less than .05 will be considered statistically significant for all the statistical tests that are conducted. Continuous variables will be tested for normality, and nonnormal values will be categorized or appropriately transformed. Given our hypotheses regarding the impact of our intervention on our key outcomes, one-sided hypothesis tests will be conducted for all our main analyses.

Our main econometric specifications will (1) estimate the intent-to-treat (ITT) effects of our family planning intervention on fertility intentions and preferences, contraceptive intentions and preferences, intention to switch methods, and other outcomes related to contraceptive behaviors by directly regressing our outcomes of interest on a binary variable indicating receipt of the intervention and (2) estimate the effect of interventions on downstream behavioral outcomes (such as women’s uptake of contraceptive options and the concordance between their revealed preference and their stated preference) using a 2-stage least squares model, in which the intermediate counseling variable will be instrumented by the treatment. The main econometric specification for estimating the ITT effect of our family planning intervention is defined as follows:

\[
y_{it} = \beta_0 + \beta_1 T_{it} + X_{it} \eta + \gamma_{00} + \epsilon_{it}
\]

where \(Y_{it}\) is the outcome variable of interest for woman \(i\) in time period \(t\), \(T_{it}\) is an indicator of assignment to 1 of the 3 treatment arms; \(X_{it}\) is a vector of individual-level covariates that are controlled for in the analysis; \(\eta\) is the individual-specific fixed effect; \(\gamma_{00}\) is the baseline level

Table 3. Outcome measures and instruments.

<table>
<thead>
<tr>
<th>Outcome or instrument</th>
<th>Baseline</th>
<th>Counseling</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude or knowledge of family planning</td>
<td>✓ b</td>
<td>✓ ✓</td>
<td>✓</td>
</tr>
<tr>
<td>Contraceptive use</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Pregnancy and fertility outcomes</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Contraceptive preferences</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Use of local health facilities</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Husband’s preferences, approval of family planning, and fertility preferences</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Women’s autonomy, empowerment, decision making, and household bargaining</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Women’s education, work, income, and employment</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Household assets</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Postpartum treatment from last birth</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Sexual and marital well-being</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Husband’s employment, work, education, and earnings</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Trust in local facilities</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Spillover of the intervention between treatment groups</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
<td>✓ ✓ ✓</td>
</tr>
</tbody>
</table>

aThe outcome was collected in the survey wave.
bNot available. The outcome was not collected in the survey wave.
of the outcome; and \( \epsilon \) is the error term. Here, the outcome variables of interest include immediate, intermediate, and long-term outcomes described in the previous sections.

We will conduct several subgroup analyses to examine how our family planning intervention effects vary across key subpopulations. Subgroups of interest include women who have or do not have their husbands’ approval for accessing contraceptive methods, women who have or have not previously used family planning methods, women whose most valued method attribute is or is not effective, working or nonworking women, higher or lower educated women, counseling sessions administered in the morning or afternoon, women who have more or less than 2 children, and counseling sessions of above or below the median length.

Finally, robustness checks (5% and 10% sample truncations and coarsening of independent variables) and falsification tests, which include placebo regression, simulation, and resampling methods, will be conducted to ascertain the strength and significance of our estimates. We will also conduct attrition-adjusted analyses to assess the extent of differential loss-to-follow-up on outcomes. In handling missing data, we will conduct a complete case analysis if the potential impact of the missing data is negligible, and we can conduct a series of sensitivity analyses, including best-worst case analyses, to generate plausible bounds of estimates across different scenarios of missingness. If we are able to assume that any missing data are missing at random, then we will conduct an ITT analysis using multiple imputation to adjust for missing data. We will then compare the ITT estimates from our multiple imputation analysis with our estimates from a complete case analysis, which would allow us to infer the extent to which missingness might impact our inference.

Sample Size and Power Calculations

Target Sample and Key Outcomes

The sample size of this study was powered to primarily identify the independent effects of tailored family counseling and male involvement on women’s contraceptive uptake and switching of methods. On the basis of the preliminary power calculations for the primary outcomes of interest, our target baseline sample consists of 700 eligible women who consent to participate in the study. Among these 700 women, 100 women were assigned to the control arm T0, and 200 women were assigned to each of the 3 intervention arms, namely, T1, T2, and T3.

Previous studies have found that one-fifth of women of reproductive age were either pregnant or had given birth to a child in the past 6 months [40]. To meet our target sample size of 700 women who would be eligible to participate in our study, we need to recruit a total of 940 women if we conservatively assumed a combined ineligibility or refusal rate of 25%.

We have powered our study to detect effects on the following outcomes: (1) contraceptive uptake and (2) intention to switch from the current contraceptive method.

Contraceptive Prevalence: Partner Invitation Interventions

Using modern contraceptive prevalence estimates for urban area of Malawi from the 2010 Malawi DHS, we assume a modern contraceptive prevalence rate of 50% among women aged 18-35 years who are currently not pregnant and who did not give birth in the 6 months before the baseline. To infer a potential effect size for our intervention, we look to evidence to a study from Zambia [28], which found that male involvement in family planning care-seeking decreased the probability of receiving family planning services by 19% over a 2-year study period. Our study differs from this study in that we do not require husbands to participate in the counseling session; instead, we leave it up to the woman to decide whether she would like to invite her husband to counseling or not. In designing our invitation in this way, a selection effect might exist, in which husbands who are more supportive of contraceptive use are more likely to comply with the intervention and participate in a joint counseling session. As a result, our male partner invitation intervention would likely lead to an increase in contraceptive prevalence.

Assuming an attrition-adjusted sample of 700 women, with 400 women assigned to partner invitation arms, T1 and T3, compared with 300 women assigned to noninvitation arms, T0 and T2, we will have 75% power to detect a percentage point increase of 10 in the contraceptive prevalence rate from 50% to 60% (\( \alpha = .05, \) two-sided). If, instead, we use a one-sided test, then we estimate that we would have 85% power to detect the same 10 percentage point increase in contraceptive prevalence.

Table 4 presents the levels of power (1 − \( \beta \)) that could be achieved for various minimum effect sizes for the modern contraceptive prevalence use, assuming a baseline contraceptive prevalence rate of 50% in both the intervention and control arms and a fixed endline sample size of 700 women.

<p>| Table 4. Power calculations for contraceptive prevalence rate, partner invitation intervention. |</p>
<table>
<thead>
<tr>
<th>Control CPR(^a), n (%)</th>
<th>Intervention CPR, n (%)</th>
<th>Significance level, ( \alpha )</th>
<th>Control sample size, n</th>
<th>Intervention sample size, n</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 (50)</td>
<td>236 (59)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.66</td>
</tr>
<tr>
<td>150 (50)</td>
<td>240 (60)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.75</td>
</tr>
<tr>
<td>150 (50)</td>
<td>244 (61)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.83</td>
</tr>
<tr>
<td>150 (50)</td>
<td>248 (62)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.89</td>
</tr>
</tbody>
</table>

\(^a\)CPR: contraceptive prevalence rate.

https://www.researchprotocols.org/2021/4/e24884

JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e24884 | p.81
(page number not for citation purposes)
Intention to Switch Methods: Partner Invitation Interventions

Our measure of a woman’s intention to switch to another contraceptive method is proxied by her answer to the question, “if you had the choice and ability to switch to another family planning method, would you choose to switch?” An intervention to encourage women to invite their husbands or male partners to counseling could have 2 potential effects on women’s stated preferences for family planning methods. On the one hand, male involvement in counseling may encourage women to change their contraceptive preferences by allowing women to more actively retain and interpret counseling messaging [15,16] and through increased communication between partners [20]. On the other hand, male involvement in counseling might discourage women from changing their stated preferences if they feel compelled to hide their true preferences from their partners [41]. On the basis of the existing evidence on male involvement, we hypothesize that male involvement in counseling is likely to increase a woman’s intention to switch methods. Assuming an attrition-adjusted sample of 700 women, with 400 women assigned to partner invitation arms, T1 and T3, compared with 300 women assigned to noninvitation arms, T0 and T2, we will have 86% power to detect a 10 percentage point increase in women’s intention to switch methods in the intervention arm from 20% to 30% (α=.05, two-sided). If, instead, we use a one-sided test, then we estimate that we would have 92% power to detect the same 10 percentage point increase in intentions to switch.

Table 5 presents the levels of power (1−β) that could be achieved for various minimum effect sizes for the likelihood of women’s intention to switch contraceptive methods, assuming a baseline switching rate of 20% in both the intervention and control arms and a fixed endline sample size of 700 women.

Table 5. Power calculations for intentions to switch methods, partner invitation intervention.

<table>
<thead>
<tr>
<th>Control likelihood, n (%)</th>
<th>Intervention likelihood, n (%)</th>
<th>Significance level, α</th>
<th>Control sample size, n</th>
<th>Intervention sample size, n</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 (20)</td>
<td>112 (28)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.69</td>
</tr>
<tr>
<td>60 (20)</td>
<td>116 (29)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.78</td>
</tr>
<tr>
<td>60 (20)</td>
<td>120 (30)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.86</td>
</tr>
<tr>
<td>60 (20)</td>
<td>124 (31)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Contraceptive Prevalence: Short, Targeted Counseling

Following a similar calculation to Table 4, we assume a modern contraceptive prevalence rate of 50% among women of reproductive age who are neither currently pregnant nor have given birth in the past 6 months at baseline. Given the lack of evidence on the role of information quantity during counseling on women’s contraceptive choices, we made some assumptions on the expected minimum detectable effect size of our targeted counseling intervention relative to the standard counseling process. Assuming an attrition-adjusted sample of 700 women, with 400 women assigned to the targeted counseling arms, T2 and T3, compared with 300 women assigned to the standard counseling arms, T0 and T1, we will have 75% power to detect a 10 percentage point increase in the modern contraceptive prevalence rate in the intervention arm from 50% to 60% (α=.05, two-sided). If, instead, we use a one-sided test, then we estimate that we would have 85% power to detect the same 10 percentage point increase in contraceptive prevalence.

Table 6 presents the levels of power (1−β) that could be achieved for various minimum effect sizes for modern contraceptive prevalence use, assuming a baseline contraceptive prevalence rate of 50% and a fixed endline sample size of 700 women.

Table 6. Power calculations for contraceptive prevalence rate, targeted counseling intervention.

<table>
<thead>
<tr>
<th>Control CPR, n (%)</th>
<th>Intervention CPR, n (%)</th>
<th>Significance level, α</th>
<th>Control sample size, n</th>
<th>Intervention sample size, n</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 (50)</td>
<td>236 (59)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.66</td>
</tr>
<tr>
<td>150 (50)</td>
<td>240 (60)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.75</td>
</tr>
<tr>
<td>150 (50)</td>
<td>244 (61)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.83</td>
</tr>
<tr>
<td>150 (50)</td>
<td>248 (62)</td>
<td>.05</td>
<td>300</td>
<td>400</td>
<td>0.89</td>
</tr>
</tbody>
</table>

αCPR: contraceptive prevalence rate.

Intention to Switch Methods: Short, Targeted Counseling

Following a similar calculation to Table 5, we assume an attrition-adjusted sample of 700 women, with 400 women assigned to the targeted counseling arms, T2 and T3, compared with 300 women assigned to the standard counseling arms, T0 and T1, and we calculate that we will have 86% power to detect a 10 percentage point increase in women’s intention to switch methods in the intervention arm from 20% to 30% (α=.05, two-sided). If, instead, we use a one-sided test, then we estimate that we would have 92% power to detect the same 10 percentage point increase in intentions to switch.

Table 7 presents the levels of power (1−β) that could be achieved for various minimum effect sizes for intention to switch contraceptive methods, assuming a baseline switching rate of 20% and a fixed endline sample size of 700 women.
Dissemination Plan

In a similar fashion to Karra and Canning [40], we formed a partnership with IPA Malawi, a US-based nonprofit research organization with operations in 42 different countries. Using rigorous research techniques, IPA works to develop and test solutions to real-world problems faced by the poor in developing countries. IPA comprises a group of leading academic researchers in development economics, behavioral economics, and psychology, based both in the United States and in developing countries. IPA Malawi has provided extensive technical and research assistance in health and development to many governmental and nongovernmental organizations in Malawi, including the MOH and the World Bank. For these reasons, we chose to partner with IPA Malawi for the study and worked with the IPA Malawi management team and a hired team of surveyors, field managers, intervention staff (family planning counselors and driver), and other support staff to conduct the fieldwork. IPA Malawi’s primary role in the study was to conduct the local field research activities, including (1) hiring, training, and management of the local field staff; (2) data collection, monitoring, and evaluation; (3) implementation of the intervention; and (4) assisting the investigators with the dissemination of results in Malawi.

In addition, we worked closely with the Malawi MOH and the Malawi RHD on dissemination and outreach activities for this study. Flip charts used for counseling were provided by the RHD for the duration of the study. Neither the MOH nor the RHD were directly involved in any of the data collection, intervention implementation activities, or other research-related activities. Finally, we collaborated with PSI Malawi, a nonprofit global health organization with programs targeting malaria, child survival, HIV, and reproductive health. PSI provided our team with family planning methods (male or female condoms) and other training resources to facilitate the administration of the counseling process.

Aggregate summary statistics and final peer-reviewed publications will be shared with participants, key partners (IPA Malawi, Dimagi, Good Health Kauma Clinic), the Malawi MOH, the Malawi RHD, and the Malawi National Statistics Office. Individual survey responses will not be shared among participants verbally, by recording, or in writing. Each interviewee’s responses will remain confidential, as per the terms of their consent to study participation.

We will produce output in peer-reviewed journals, working papers, and policy briefs that are accessible to academics, policy makers, and practitioners and contribute to policy changes in this area. Aggregate results and final publications will be disseminated to the community and local institutions where the research was conducted (University of Malawi, among others). The research team will also present intervention findings at local and national venues, including annual meetings of professional organizations, community gatherings, and meetings with local service providers. These meetings and conferences will be valuable opportunities for us to see the rigor of work that our contemporaries are working toward and will allow us to contribute our own progress in the field of economic outcomes of reproductive health.

Our work will also be effectively disseminated to practitioners via local partnerships with FHI360, BLM, PSI Malawi, the World Bank, the RHD, and the MOH in Malawi. We have worked with these organizations during the study design phase to ensure that the interventions are appropriate for the country setting. Upon the completion of the intervention, the local partners can leverage the study findings to expand or tailor their services to help women achieve their family planning goals within a specific country context. We will also share our descriptive and analytical findings with members of the community who are engaged in advocacy efforts.

Our dissemination efforts are engrained in our interventions from the outset. We will tailor our information packs and materials for women in Malawi based on knowledge gained from preliminary studies of the family planning environment. Consulting practitioners and local agencies will develop the information packs and other counseling components of the intervention with wording that has integrity and clarity, that uses language as commonly used by local women, that is pitched at a layperson level, and that is sensitive to local culture and practices. This dissemination of information as a part of the intervention will help us learn how to further disseminate the results of our research to the women in the communities we are working in.

Protocol Amendments and Modifications

All protocol amendments and modifications were submitted to both the BU IRB and the Malawi NHSRC for approval. Before the implementation of any modified protocols, the approval of the modification had to be received by both ethics committees. Any determination of the extent to which modifications were communicated to the study participants was made by both ethics committees at the time of review.

Data Confidentiality

Our data confidentiality protocols follow those established by Karra and Canning [40]. To effectively monitor participants over the study period, identifiable data on participants’ demographic backgrounds and personal contact information (household addresses, phone numbers, emails, and GPS
locations) were collected. In addition, photographs of participants and of the households were taken to facilitate identification at the follow-up stage. All identifiable data collected from surveys (both baseline and follow-up surveys) and from the intervention were administered in an electronic computer-assisted personal interview format using the CommCare survey management system. Electronic survey data were collected by interviewers on Android-based tablets, and data were securely transferred from the Android tablets onto a CommCare-supported secure cloud server at the end of each working day. All Android tablets were used for data collection only, and tablet settings were adjusted so that field staff was blocked from accessing apps that were not relevant for data collection (eg, internet browsing, social media, and email). The CommCare cloud server was Health Insurance Portability and Accountability Act (HIPAA)–compliant and met all the necessary security requirements for storing level 4 identifiable data. Once the data were securely transferred to the cloud server, the survey record on the Android tablet was immediately erased. A technical overview of the CommCare system, including descriptions of the data transfer process and the HIPAA-compliant storage system, can be found on the Dimagi website [42]. An electronic version of the CommCare Terms of Use or End User License Agreement can also be found on the Dimagi website [43].

All data uploaded to the CommCare cloud server were encrypted and password-protected in accordance with level 4 data security and storage regulations. For each collected data case, which comprised a woman’s (and her husband’s) data records, all personal identifiable data were separated from the other nonidentifiable data. The deidentified data were uploaded to an encrypted password-protected File Transfer Protocol (FTP) site on a daily or weekly basis and circulated to the project PIs for analysis purposes. The identified data were stored on the CommCare secure encrypted server and could only be accessed for the purpose of revisiting the households at the 1-month follow-up period. After the study ended, the IPA Malawi research site maintained the identified data in an encrypted file on a secure server, and only deidentified data sets remained available for analysis purposes after the end of the study.

Every effort will be made to ensure that participation in this study and all records about participation remain confidential. As previously stated, all confidential identifiable data were secured by trained study personnel upon collection. Data were collected by trained staff and fully deidentified as soon as possible. We will work with Dimagi to set up a data management system that meets the following requirements:

1. Raw electronic survey data will immediately be transferred to the CommCare secure cloud server once it has been collected on the Android-based tablets. Following the transfer, the data from the Android tablets will be automatically erased.

2. All identifying information will be separated from the raw electronic survey data immediately after collection and secure transfer to the cloud server, and a unique computer-generated ID number, which is created when a respondent is registered in the study database, will be assigned to each case. Coded, deidentified data files will be stored separately from the code list and the identified data files. Following the completion of all field research activities, all respondents will only be referred to by their ID number in the deidentified data set. Only the PIs will have access to the linkages to the underlying identifiable files. Identifiable electronic data will be encrypted, password-protected, and securely stored on the CommCare-protected cloud server, and one copy of the data will be stored on a password-protected target computer.

3. Restrictions will be placed on nonauthorized users from accessing certain data or features by assigning them permission levels. This includes restricting access to any identifying data that would violate HIPAA or other privacy standards. Each study team member will be assigned 1 of 3 permission levels, which will provide them with varying levels of access, from no access (level 0) to full access (level 2).

Identifiable hard-copy data, including signed consent forms, were stored in locked cabinets in access-limited rooms at the IPA Malawi office. All study computers that are used for descriptive analysis of the deidentified data are password-protected, and only study staff who are cleared to view the data will have the password. All electronic data, both on the CommCare secure cloud server and on any study computer, will be encrypted and password-protected. This information will be accessible only to the research team.

All staff members of the study were required to sign a data confidentiality agreement. The data were stored in a relational database. Usernames and passwords are required to access the data. A security policy is used to ensure that the passwords are updated on a regular basis.

Data sharing of deidentified data between the immediate research team will be conducted in person—a USB key will be used to transfer the deidentified data from one secure hard drive to the next, and data will then be deleted from the USB key. The USB key will be used only for storing and transferring research material between the research team members mentioned above, and it will not be used for storing or transferring other files that are unrelated to the study. Four sets of deidentified data will be stored, one for each of the immediate research team members.

Deidentified data will be transmitted from Malawi to BU via secure file transfer (through the BU system). Colleagues in Malawi will be offered guest access to transfer the data. All hard-copy data and electronic data will be retained for 7 years after study closure, after which it will be destroyed (shredding hard copies and permanently deleting all electronic files).

**Data Safety and Monitoring**

The PI, MK, will take the overall responsibility for the safety, monitoring, and review of the data. He and KZ will oversee the weekly review of all data collected in the study and were present during the administration of the baseline, intervention, and follow-up surveys. He and the coinvestigators of the study will ensure that the data are treated as confidential and stored in a secure location, as detailed above.
For the proposed research, the local project manager at IPA Malawi; the local coinvestigators, Bagrey Ngwira and Abiba Longwe; and the PI, MK, will review adverse events and protocol deviations. This information will then be provided to the institutional review boards at the BU and the NHSRC in Lilongwe. Unanticipated adverse events and protocol deviations will be immediately reported to both the BU IRB and the NHSRC in writing within 5 business days. As per our stated reporting protocols, we will also inform local community leaders of any adverse events related to participant safety, domestic violence, and abuse within the household, and we will refer participants and other members of the household to their nearest victim support unit and to the Department of Social Welfare at the Lilongwe District Council Office. In addition to reporting any events from the baseline and follow-up surveys, we will make quarterly reports on consultations and reimbursements for family planning services for women.

We do not anticipate that there will be any research-related injuries. Bagrey Ngwira, Abiba Longwe, and the IPA Malawi project manager trained the field staff (surveyors, field managers, and interventionists) in first aid and will be on call via mobile phone throughout the entire duration of the study. The field team was trained to recognize basic signs of physical and psychological injuries and were instructed to immediately report any injuries that were incurred during the interview with the IPA Malawi project manager. With the help of Bagrey Ngwira, Abiba Longwe, and the IPA Malawi project manager, we established a link with a local primary clinic in Lilongwe and will refer respondents to this clinic if they experience research-related injuries. We will identify the emergency care facility or hospital that is nearest to the interview location before each interview, and we will refer respondents to this facility in case of a medical emergency.

**Regulatory Compliance**

MK, KZ, Bagrey Ngwira, and the IPA Malawi project manager will be the immediate supervisors of the study staff in the field. They will communicate with the study staff on a weekly basis and will ensure that the study protocol and IRB regulations are being followed. On-site supervision will allow the management team to provide support for staff as well as quality assurance and confidentiality of study data. In addition, the local team and Boston-based team will be in regular contact via email in the interim to discuss the progress of the study protocol procedures. MK and KZ will travel regularly to Malawi over the course of the study period and particularly during the data collection phases to monitor field activities. All regulatory documentation will be maintained for 7 years after IRB study closure.

**Authorship Eligibility Guidelines**

A publication committee comprising MK, the overall PI, and KZ has been established to address and decide on all matters related to access to project data and publications using such data. All guidelines for data access and publications are outlined in the publications committee terms of reference document (available upon request).

**Availability of Data and Material**

Following our own use and analysis of the data (a minimum 1-year period), we hope to open access to deidentified baseline and follow-up survey data at no cost to authorized users. Only deidentified data will be available for download via a secure website, through which authorized users can download deidentified survey data files for legitimate academic research. To access the data, prospective users must first register on the secure website and must then create a new research project request. The request must include a project title and a description of the analysis that the user proposes to perform with the data. The requested data should only be used for research or study purposes. To request the same data for another purpose, a new research project request needs to be submitted. Requests for data access will then be reviewed by the PI, who can then grant or deny access to the user. All publications that users produce from the data set must appropriately acknowledge the data source and project from which the data were collected. Once downloaded, the data sets must not be passed on to other researchers without written consent from the PI. All reports and publications based on the requested data must be sent via email to the BCS in a portable document format or as a printed hard copy.

**Results**

**Recruitment, Study Sample, and Randomization**

Field activities for the baseline survey began with field staff hires, training, and piloting of the survey instrument in June 2019 and continued through September 2019. During the 3-month baseline survey period, 1122 women were approached and screened using the eligibility criteria. On the basis of the eligibility screening, of the 1122 women, 823 (73.4%) were found to be eligible to participate in the study. Of the 823 eligible women, 801 agreed to go through the consent form with the enumerator, and 785 (95.3% of the eligible sample) further consented to participate and were subsequently enrolled in the study. Of these 785 women, 782 (99.6%) completed the baseline survey and were eligible to be randomized into 1 of the 3 intervention groups or the control group. From this baseline sample, 223 women were randomly assigned to treatment (T1) with long counseling and husband invitations, 225 women were assigned to treatment (T2) with short counseling and no husband invitations, 228 women were assigned to treatment (T3) with short counseling and husband invitations, and 108 women were assigned to treatment (T0) with long counseling and no husband invitations. Among the 782 women in the final sample, 41 were interviewed as a part of a preliminary pilot study to test the feasibility of the survey instruments and implementation of the intervention. As part of the intervention rollout, these 41 respondents were also randomized into the T1 (n=10), T2 (n=13), T3 (n=12), and control T0 (n=6) arms. The final analytic sample for the baseline survey comprised 782 eligible women, of which 223 were assigned to T1, 225 were assigned to T2, 228 were assigned to T3, and 108 were assigned to T0 (control). The experimental framework is illustrated in Figure 1.
**Intervention Activities**

The rollout of the counseling intervention began shortly after randomization in September 2019. Four family planning counselors (registered nurses and midwives with previous counseling experience in family planning) were identified in August 2019 and were trained through September 2019 to administer the range of counseling services over a 3-month intervention period. Counseling in the intervention groups began in September 2019.

In addition to hiring 4 counselors, the study management team hired and trained a licensed taxi driver in March 2019 to assist with the implementation of the transportation component of the intervention. In March 2019, the management team also identified an obstetrician at the Kamuzu College of Medicine as a medical doctor on call. The obstetrician was asked to be responsible for (1) answering any calls from clients; (2) providing consultation services and information on the desired family planning methods for any woman who visited the clinic during the service period; (3) providing any support or consultation services over the phone; and (4) referring any client who may be experiencing health concerns, particularly those related to their use of family planning, to the management team for follow-up.

Counseling activities with women in the intervention group concluded in December 2019; however, other intervention activities (providing transportation to women to visit the Kauma Clinic for services and providing financial reimbursements to women for any family planning services that they obtain) continued until February 2020.

**Follow-Up Surveys**

Follow-up surveys began in November, 2 months after the conclusion of the baseline survey. Field activities for the follow-up surveys began with hiring field staff, training, and piloting the follow-up survey instruments in November 2019, which continued through February 2020. During the 4-month follow-up survey period from November 2019 to February 2020, 767 women (including 727 women from the main study and an additional 40 women from the pilot phase of the study, but not including the 15 women who withdrew from the study before the start of the follow-up survey) were reached. Of the 767 women who were reached at the follow-up stage, 731 (95% of women who were eligible for follow-up) successfully completed the follow-up survey. In total, 51 women who were interviewed at baseline are estimated to be lost to follow-up.

**Analysis**

All study-related field activities for the follow-up survey were completed on February 15, 2020. The cleaning of baseline, precounseling, postcounseling, and follow-up survey data are ongoing. A complete, cleaned survey wave includes the following: a recoded and indexed data set, a data codebook, a recode map and variable list, a final survey questionnaire, a final report and user manual for the survey wave, and data analysis files and templates. Analyses of the primary outcomes are ongoing and are expected to be completed by 2021.

**Discussion**

**Principal Findings**

The primary objective of this study is to investigate the extent to which user-centered approaches to counseling, by exploiting choice architectural elements around family planning and promoting male involvement, can affect a woman’s immediate stated preferences and realized preferences in a context where her preferences may be sensitive to a range of behavioral biases. Our main outcomes of interest include women’s stated preferred method of contraception, intention to switch to other contraception methods, uptake of their ideal contraceptive method, and realization of their ideal contraceptive preferences. In addition, by observing whether women will seek family planning following their counseling visit, we will examine the short-run stability of their stated preferences and the extent to which these stated preferences are subsequently realized in the face of barriers to use. Data examining the realization of women’s eventual contraceptive preferences (eg, contraceptive uptake, eventual method choice) were collected during clinic visits and at (phone or home) follow-ups across the four arms.

**Study Limitations**

Our study has several limitations. Although the study is powered to independently test the effects of 2 user-centered counseling procedures on women’s stated and realized preferences, the sample size of this study may be too small to allow for the examination of interaction effects across treatment arms. As an exploratory exercise, we will examine these effects and present them as supplementary findings. Moreover, although our study can infer the manner in which women realize their preferences in response to counseling, it is not as straightforward to be able to disentangle a woman’s true individual preferences for family planning from her stated preferences in the presence of her partner, who is likely to influence her eventually realized choice (or lack of choice) of method. To this end, additional research is warranted to better understand how women’s true preferences are expressed and can be documented, even in cases where their partner is present. Similarly, it would be equally necessary to further explore the trade-off that women face between (1) making independent decisions to reflect their individual preferences and (2) incorporating their husband’s/family’s preferences to make jointly/socially better-off, but not necessarily individually better-off, decisions.

**Conclusions**

Our study findings seek to inform women and communities of the local family planning environment. In addition, our findings will inform both clients and service providers of the role of user-centered, tailored family planning counseling in improving concordance between a woman’s preferences and her contraceptive use, which in turn would help her realize her family planning goals. The results may also help the MOH and other service providers to develop family planning programs that internalize the preferences of women and their partners for contraceptive methods, thereby achieving the goal of improving reproductive health outcomes in Malawi. More generally, our findings may also demonstrate to policy makers that the benefits of tailoring family planning services to women’s preferences
are likely to extend beyond the health domain by improving empowerment outcomes. Such findings can be used to develop policies, programs, and interventions that aim to improve the health and well-being of women, couples, and households.

Acknowledgments
The authors are grateful to Patrick Baxter, Bagrey Ngwira, and Abiba Longwe-Ngwira. This research makes use of original data collected by MK and KZ with support from IPA in Malawi. Field implementation would not have been possible without the support of the Malawi MOH and RHD; PSI Malawi; Faison Mussa; the Good Health Kauma Clinic; Patrick Baxter; and the entire Malawi Behavioral Biases Study team, which comprised 5 enumerators and 4 family planning counselors over an 8-month study period. The authors would also like to thank Amelia Dangerfield and Leah Eyob for their research assistance. Finally, the authors would like to acknowledge the dedication and support of Violet Chitsulo and Macdonald Salamu for their coordination and support in the data collection phase. All errors are our own. Funding support for the main study was provided by the William and Flora Hewlett Foundation (Grant 2018-8083). KZ received travel support from the Manuel Abdala Travel Grant and the Institute of Economic Development Summer Research Grant provided by the BU Economics Department. The funding sources had no role in the design and implementation of this study and will have no role in the data analyses, interpretation of results, or the dissemination of findings.

Authors' Contributions
All authors contributed equally to the design and development of the study and to the production of the study protocol. MK and KZ contributed to various sections and drafted the manuscript, the development and revision of the study instruments, power calculations, sampling, implementation plans, and plans for data analysis. All authors read and approved the manuscript for publication.

Conflicts of Interest
External funding for the study was provided through two grants from the William and Flora Hewlett Foundation, Menlo Park, California, United States. The funder had no role in the design and implementation of this study and will have no role in the data analyses, interpretations of results, or dissemination of findings. In addition, the study protocol did not undergo peer review by the funding body. Finally, all study investigators have read and understood the BU IRB and NHSRC policies on the declaration of competing interests. The authors declare that they have no conflict of interest, financial or otherwise, other than the normal scholarly gains from taking part in this study.

Multimedia Appendix 1
Survey instruments, recruitment scripts, and consent forms. [PDF File (Adobe PDF File), 2295 KB - resprot_v10i4e24884_app1.pdf ]

Multimedia Appendix 2
Family planning flip charts. [PDF File (Adobe PDF File), 22873 KB - resprot_v10i4e24884_app2.pdf ]

References
4. Sennott C, Yeatman S. Stability and change in fertility preferences among young women in Malawi. URL: https://doi.org/10.1363/3803412 [accessed 2021-03-02]
9. Schlag P. Nudge, choice architecture, and libertarian paternalism. URL: https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1278&context=mlr [accessed 2021-03-02]


21. McCarthy AS. His and her fertility preferences: an experimental evaluation of asymmetric information in family planning. 2015. URL: https://www.semanticscholar.org/paper/His-and-Her-Fertility-Preferences%3A-An-Experimental-McCarthy/fe71f3631f2d0db1b98c243efc774f821e045 [accessed 2021-03-02]


40. Karra M, Canning D. The effect of improved access to family planning on postpartum women: protocol for a randomized controlled trial. JMIR Res Protoc 2020 Aug 14;9(8):- [FREE Full text] [doi: 10.2196/16697] [Medline: 32795987]


Abbreviations

- BCS: Balanced Counseling Strategy
- BLM: Banja La Mtsogolo
- BU: Boston University
- DHS: Demographics and Health Survey
- HIPAA: Health Insurance Portability and Accountability Act
- IPA: Innovations for Poverty Action
- IRB: institutional review board
- ITT: intent-to-treat
- MDHS: Malawi Demographics and Health Survey
- MOH: Ministry of Health
- NHSRC: National Health Sciences Research Committee
- PI: principal investigator
- PSI: Population Services International
- RHD: Reproductive Health Directorate

©Mahesh Karra, Kexin Zhang. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 05.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Development and Implementation of a Mobile Tool for High-Risk Pregnant Women to Deliver Effective Caregiving for Neonatal Abstinence Syndrome: Protocol for a Mixed Methods Study

Ekaterina Burduli1,2, PhD; Hendrée E Jones3, PhD; Olivia Brooks1, MS; Celestina Barbosa-Leiker1, PhD; Ron Kim Johnson4, MBA; John Roll5, PhD; Sterling Marshall McPherson2,5, PhD

1College of Nursing, Washington State University, Spokane, WA, United States
2Analytics and PsychoPharmacology Laboratory, Washington State University, Spokane, WA, United States
3UNC Horizons, Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
4Managed Health Connections, LLC, Spokane, WA, United States
5Elson S. Floyd College of Medicine, Washington State University, Spokane, WA, United States

Corresponding Author:
Ekaterina Burduli, PhD
College of Nursing
Washington State University
412 E Spokane Falls Blvd
Spokane, WA, 99202
United States
Phone: 1 509 324 7368
Email: eburduli@wsu.edu

Abstract

Background: The United States continues to experience an alarming rise in opioid use that includes women who become pregnant and related neonatal abstinence syndrome (NAS) in newborns. Most newborns experiencing NAS require nonpharmacological care, which entails, most importantly, maternal involvement with the newborn. To facilitate positive maternal-newborn interactions, mothers need to learn effective caregiving NAS strategies when they are pregnant; however, an enormous gap exists in the early education of mothers on the symptoms and progression of NAS, partly because no education, training, or other interventions exist to prepare future mothers for the challenges of caring for their newborns at risk for NAS.

Objective: In this paper, we describe a mixed methods, multistage study to adapt an existing mobile NAS tool for high-risk pregnant women and assess its usability, acceptability, and feasibility in a small randomized controlled trial.

Methods: Stage 1 will include 20 semistructured interviews with a panel of neonatology experts, NAS care providers, and mothers with experience caring for NAS-affected newborns to gather their recommendations on the management of NAS and explore their perspectives on the care of these newborns. The findings will guide the adaptation of existing mobile NAS tools for high-risk pregnant women. In stage 2, we will test the usability, acceptability, and feasibility of the adapted mobile tool via surveys with 10 pregnant women receiving opioid agonist therapy (OAT). Finally, in stage 3, we will randomize 30 high-risk pregnant women receiving OAT to either receive the adapted mobile NAS caregiving tool or usual care. We will compare these women on primary outcomes—maternal drug relapse and OAT continuation—and secondary outcomes—maternal-newborn bonding; length of newborn hospital stays; readmission rates; breastfeeding initiation and duration; and postpartum depression and anxiety at 4, 8, and 12 weeks postpartum.

Results: This project was funded in July 2020 and approved by the institutional review board in April 2020. Data collection for stage 1 began in December 2020, and as of January 2021, we completed 18 semistructured interviews (10 with NAS providers and 8 with perinatal women receiving OAT). Common themes from all interviews will be analyzed in spring 2021 to inform the adaptation of the NAS caregiving tool. The results from stage 1 are expected to be published in summer 2021. Stage 2 data collection will commence in fall 2021.

Conclusions: The findings of this study have the potential to improve NAS care and maternal-newborn outcomes and lead to commercialized product development. If effective, our new tool will be well suited to tailoring for other high-risk perinatal women with substance use disorders.
Introduction

Background
The United States is experiencing an alarming rise in opioid use during pregnancy, diagnosis of opioid use disorder (OUD) in pregnancy, and related neonatal abstinence syndrome (NAS) in newborns [1,2]. The percentage of pregnant women reporting opioid misuse increased from 2% to 28% in the United States from 2000 to 2014, and the proportion of pregnant women who entered substance use treatment and reported prescription opioids as their primary substance increased from 16.9% in 1996 to 41.6% in 2014 [3]. NAS is a spectrum of symptoms or signals of substance withdrawal in newborns whose mothers used illicit opioids or were being treated with opioid agonist therapy (OAT; ie, methadone or buprenorphine) during pregnancy [4,5]. NAS rates have dramatically increased from approximately 1.5 to 8.0 per 1000 hospital births from 2004 to 2014 [5-8], which translates to one infant experiencing opioid withdrawal every 15 minutes in the United States [1]. There is evidence that this rate continues to increase as the US Pediatric Health Information System reported an incidence of 20 per 1000 live births in 2016 [9]. NAS signs include tremors, increased irritability, fever, sweating, extreme weight loss, excessive sucking, and inability to sleep [4,10,11]. NAS is an enormously costly public health issue, costing the United States US $563 million in an aggregate associated hospital charge in 2014, largely because of costs associated with lengthy hospital stays, neonatal intensive care unit (NICU) admissions, and higher risk of hospital readmission, and other costly, downstream health costs associated with depression, anxiety, and other difficulties that mothers encounter [1,6,12,13].

Mothers are often unprepared for what they are about to experience before having a baby while using either illicit or licit opioids. Treatment of NAS in the postpartum period is critical in shaping both maternal and newborn outcomes. At least half [14] of the newborns experiencing NAS require nonpharmacological care, which is best delivered by the newborn’s mother or close caregiver [4,15-17]. Up to 95% of NICUs offer some form of nonpharmacological intervention, and it is advised that all infants with in utero exposure to opioids be administered nonpharmacological treatment [14]. Evidence-based, nonpharmacological care for NAS includes rooming together with the mother postdelivery and modification of the environment to support maternal-newborn attachment that promotes a soothing environment for the infant. Specific actions include low-stimulating environments (reduced noise and light), frequent and small feedings, swaddling, promotion of breastfeeding, and continual contact and soothing from the caregiver [15,17-20]. These are all doable now by mothers and other health care providers; however, the missing link is having ready access to an educational platform to promote these strategies, which is precisely what this study is designed to do.

Facilitating successful postpartum maternal-newborn involvement is critical in lowering NICU admissions and hospital readmissions, boosting breastfeeding, shortening hospital stays, decreasing the need for pharmacological intervention for the newborn, reducing the risk of maternal OAT discontinuation, and facilitating maternal-child bonding [13,15,21-24]. However, the challenging NAS signs can hinder maternal-newborn bonding, particularly for women who have substance use disorders (SUDs) who may have difficulty responding to newborns’ cues [15]. To maximize positive maternal-newborn interactions, mothers need to learn effective NAS caregiving strategies when they are pregnant to optimize their preparedness; however, little to no preparedness strategies exist for at-risk pregnant women.

The American College of Obstetricians and Gynecologists' Alliance for Innovation on Maternal Health Program recently developed a maternal safety bundle on obstetric care of women with OUDs, which emphasizes the need to provide caregivers education regarding NAS and newborn care [19]. However, pregnant women in OAT currently receive virtually zero education or consultation for what to expect when they give birth to a baby who is at risk of experiencing NAS. One such education strategy would incorporate caregiving skills based on the promising and recently established Eat, Sleep, and Console (ESC) model that stresses newborn sleep, feeding, and a low-stimulating environment [20]. We previously developed and implemented a mobile instructional tool designed to streamline the assessment of NAS for providers in the NICU. This interactive tool includes evidence-based education modules covering topics such as NAS epidemiology, symptoms, nonpharmacological treatment, and transition to follow-up care [25].

Objectives
As the long-term caregivers of newborns experiencing NAS, mothers are essential to the successful, nonpharmacological treatment of NAS-affected babies. By adapting our instructional NAS tool, we can address this critical gap and potentially improve outcomes for NAS-affected newborns and their mothers. This study describes a mixed methods, multi-stage study (see Figure 1 for study stage and timeline) to adapt an existing mobile NAS tool for high-risk pregnant women and assess its usability, acceptability, and feasibility in a small randomized controlled trial.
Methods

Overview
This study will be carried out in 3 stages. In stage 1, we will conduct semistructured interviews with a panel of NAS care experts and mothers with NAS-affected newborns until saturation is reached (expected n=20) to document their perspectives and gather their recommendations on the care of newborns with NAS. Although findings from stage 1 will ultimately guide the adaptation of the existing mobile NAS tool for high-risk pregnant women, the skills training in the adapted tool will be largely based on elements of ESC. In stage 2, we will test the usability, acceptability, and feasibility of the adapted mobile tool via surveys with 10 pregnant women receiving OAT at 2 facilities in Washington: one that provides methadone as its pharmacotherapy and another that primarily treats patients receiving buprenorphine as their pharmacotherapy. In stage 3, we will randomize 30 high-risk pregnant women seen at these facilities to either receive the adapted mobile NAS instructional tool or usual care. We will compare these participants on numerous maternal and newborn health outcomes at 4, 8, and 12 weeks postpartum, including primary outcomes—maternal OAT continuation and maternal drug relapse—and secondary outcomes—length of newborn hospital stay, readmission rates, maternal-newborn bonding, breastfeeding initiation and duration, birth satisfaction, and postpartum depression and anxiety.

Adapting the Mobile NAS Tool for High-Risk Pregnant Women
All videos in the existing modules for NICU providers have a script, with different content covered in each video. We will modify these modules (eg, titles, scripts, and content reading levels) to increase their relevance for high-risk pregnant women and for nonexperts. We expect that the adaption process will require (1) the removal of existing educational modules that are not appropriate for at-risk pregnant women (eg, Epidemiology and Pathophysiology of NAS), (2) reduction of video lengths, and (3) development of modules based on qualitative findings that are appropriate for high-risk pregnant women (eg, ESC-based caregiving skills for mothers, postpartum maternal resources, preparing for hospital length of stay, and preparing for possibly experiencing health care stigma). For example, one existing module is titled Epidemiology and Pathophysiology of NAS, which provides considerable detail and medical terminology. We expect that this module will need to be substantially reduced and only include a brief overview of this topic in a way that is appropriate for at-risk pregnant women. We also expect that the adapted tool will include brand-new educational modules and accompanying videos that visually demonstrate ESC-based caregiving skills that are known to soothe a NAS-affected newborn and prevent further complications. We expect that the total number of modules for perinatal women will not exceed 6.

After this initial process of modification, the modules will be uploaded to a private account on YouTube, and all participating individuals (ie, NAS providers and mothers with NAS-affected babies) will be given free web access (via a YouTube link) to the existing modules and will have the opportunity to review the existing content before providing feedback on further necessary changes and modifications. They will be asked to provide their opinion on the relevance of a new audience (pregnant women in treatment for OUD). Once semistructured interviews are completed and common themes are analyzed using qualitative description methodology, we will schedule a mini conference with the participants of this study stage (ie, NAS care experts and mothers with NAS-affected babies) to discuss the interpretation of the qualitative findings and double check for consistency that the planned adaptation of the mobile NAS tool matches what the participants tried to convey.

Although findings will ultimately guide the adaptation of the existing mobile NAS tool for high-risk pregnant women, the skills training in the adapted tool will be largely based on elements of ESC. Therefore, we anticipate that the NAS mobile tool intervention will incorporate nonintrusive caregiving skills and strategies that encompass providing a low-stimulating environment (eg, dimmed light and low noise), swaddling, continuous comfort and contact with the caregiver, skin-to-skin contact, frequent feeding, and novel components identified in the interviews.
Stage 1: Adaptation of the Mobile NAS Instructional Tool for Pregnant Women at Risk for Having Newborns With NAS

Setting and Data Collection

We will recruit 10 pregnant and postpartum women from 2 locations: a regional health district’s opioid treatment program and local recovery center to participate in the semistructured interviews. Recruitment materials with contact information for study staff and a link to a brief information statement that describes the screening process and eligibility requirements will be distributed to the 2 programs via email and presented at staff meetings to recruit the 10 pregnant and/or postpartum participants. At least 10 NAS care experts (broadly defined as any person providing NAS services such as, but not limited to, nurses, social workers, obstetricians and gynecologists, and neonatologists) will also be identified and contacted from existing partnerships with local and national hospitals. After a verbal or PowerPoint description of the adapted module topics and content is provided and participants have a chance to ask questions, we will ask participants to provide feedback on the module topics and content (eg, “What do you think of the modules?” “Do you think you could use it with relative ease? What are any challenges to using it?” “Are any topics missing from the modules?” “Do you think this would help you care for a baby with Neonatal Abstinence Syndrome after birth? If yes, how? If no, what could improve it?” “What do you think we should add to this study?” “Any additional thoughts and suggestions about the modules?”). Each participant who completes the interview will receive a US $25 Amazon e-gift card for their time.

Inclusion Criteria

Eligibility criteria were selected to allow for an accurate assessment of challenges to NAS care and necessary evidence-based skills to care for NAS-affected newborns, while maximizing generalizability of results. Inclusion criteria for NAS care experts are (1) NAS providers with significant expertise in NAS care, (2) age >18 years, and (3) ability to speak and understand English. The inclusion criteria for pregnant and postpartum participants are (1) a pregnant woman currently in OAT for OUD or a postpartum woman who has experience caring for NAS-affected newborns, (2) age >18 years, and (3) ability to speak and understand English. Participants not meeting the eligibility criteria will be excluded.

Data Analysis

All interviews will be digitally audio recorded and then transcribed verbatim by a professional transcriptionist. The deidentified interview transcriptions will be uploaded into NVivo version 11, a software program produced by QSR International for qualitative research. The qualitative description methodology described by Schreier [26] will be used to analyze the data. A qualitative description methodology is used when the goal is to summarize descriptions of events or experiences in a way that depicts the perspectives of the participants [27,28]. Common themes will be identified in the data (transcripts) to provide definitions and details of the most prominent ideas provided by participants’ responses. The qualitative description methodology will involve combing concept-driven and data-driven analysis approaches to the text; reliability will be addressed by the process of having each researcher (ie, principal investigator, research coordinator, and research team) initially review and analyze the data before each meeting and then comparing consistency of agreement between the researchers, and validity will be addressed by considering the applicability of the themes when compared with participants’ responses. An audit trail will be kept throughout the analysis process to document decisions and next steps.

Stage 2: Usability, Acceptability, and Feasibility of the Adapted NAS Tool

Data Collection and Outcomes

We will recruit 10 pregnant women receiving OAT from 2 locations: a regional health district’s opioid treatment program and a local recovery center to participate in the survey, as described earlier in stage 1. Consented participants will be asked to review the adapted content (via a private YouTube link) before completing the surveys. Once they review the adapted modules, they will be directed to complete a Research Electronic Data Capture survey that will assess several measures. Acceptability will be examined by the 8-item Client Satisfaction Questionnaire (CSQ-8) to rate the overall satisfaction with the adapted NAS tool [29]. Usability will be assessed using the 10-item Systems Usability Scale (SUS), in which they are asked to answer questions about a mobile app (eg, “I think that I would need the support of a technical person to be able to use this app”) [30,31]. To assess feasibility, a measure of utility will be assessed via several questions (eg, “To what extent do you expect to be able to incorporate the NAS caregiving tool in your daily activities during pregnancy and postpartum?”), tracking how many times participants referred to specific modules within the mobile tool, and open-ended questions that ask participants to comment on the feasibility of this tool along with overall impressions and comments on the tool (survey details are given in Tables 1 and 2). Participants will receive a US $25 e-gift card for their time.
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Baseline (during the third trimester)</th>
<th>4, 8, and 12 weeks postpartum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>✓&lt;sup&gt;a&lt;/sup&gt;</td>
<td>N/A&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Education level</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td>Marital status</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td>Time in OAT&lt;sup&gt;c&lt;/sup&gt; program</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td>Number and age of living children</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td>Employment</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Baseline characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Addiction Severity Index-Lite</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td>Maternal Antenatal Attachment Scale</td>
<td>✓</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Outcome measures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Addiction Severity Index-Lite</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Maternal Postpartum Attachment Scale</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Birth Satisfaction Scale-Revised</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Patient Health Questionnaire-9</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Parenting Stress Index</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Client Satisfaction Questionnaire-8</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>System Usability Scale</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Length of hospital stay for newborn</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Newborn hospital readmission</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Breastfeeding initiation and duration</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Maternal OAT continuation and relapse</td>
<td>N/A</td>
<td>✓</td>
</tr>
<tr>
<td>Frequency of NAS&lt;sup&gt;d&lt;/sup&gt; tool use (weekly)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

<sup>a</sup>Data collected.

<sup>b</sup>N/A: not applicable.

<sup>c</sup>OAT: opioid agonist therapy.

<sup>d</sup>NAS: neonatal abstinence syndrome.
Table 2. Outcomes and description of measures.

<table>
<thead>
<tr>
<th>Study stage and outcome</th>
<th>Study design</th>
<th>Description of measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 1: NAS® tool adaptation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tool adaptation</td>
<td>Semistructured interviews</td>
<td>Semistructured interviews with a panel of NAS care experts and mothers with NAS-affected newborns until we reach saturation (expected n=10) to document their perspectives and gather their recommendations on the care of newborns with NAS.</td>
</tr>
<tr>
<td><strong>Stage 2: Usability, acceptability, and feasibility of the adapted NAS tool</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usability</td>
<td>Survey</td>
<td>Participants will complete a 10-item SUS, in which they are asked to answer questions about a mobile app (eg, “I think that I would need the support of a technical person to be able to use this app”) using a 5-point Likert scale, ranging from strongly disagree to strongly agree. The SUS is currently the industry standard for evaluation of a wide variety of products and services such as software, mobile devices, websites, and apps [30,31]. The SUS has been shown to successfully differentiate between usable and unusable systems, and it can be used in small samples with reliable results [30,31].</td>
</tr>
<tr>
<td>Acceptability</td>
<td>Survey</td>
<td>Acceptability will be examined by the CSQ-8 to rate the overall satisfaction with the adapted NAS tool [29]. CSQ-8 is a validated questionnaire measuring satisfaction with health services (eg, “How would you rate the quality of care you received?”) that has been translated in more than 30 languages [29,32,33]. Possible total scores on CSQ-8 range from 8 to 32, with higher scores (&gt;23) indicating greater satisfaction with health services.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>Survey</td>
<td>To assess feasibility, a measure of utility will be assessed (eg, “To what extent do you expect to be able to incorporate the NAS caregiving tool in your daily activities during pregnancy and postpartum?”), tracking how many times participants referred to specific modules within the mobile tool and open-ended questions that ask participants to comment on feasibility of this tool along with overall impressions and comments.</td>
</tr>
<tr>
<td><strong>Stage 3: Clinical trial comparing NAS tool with usual care</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics and baseline characteristics, such as age, education, marital status, time in the OAT program, number and age of living children, employment, and ASI-Lite, will be collected at the start of the study to describe the sample and to serve as control variables when comparing groups on the outcomes.</td>
<td>Randomized clinical trial</td>
<td>Demographics and baseline characteristics, such as age, education, marital status, time in the OAT program, number and age of living children, employment, and ASI-Lite, will be collected at the start of the study to describe the sample and to serve as control variables when comparing groups on the outcomes.</td>
</tr>
<tr>
<td>Maternal drug use and relapse</td>
<td>Randomized clinical trial</td>
<td>Maternal drug use and relapse will be assessed via ASI-Lite, a standardized semistructured clinical interview that offers clinical information and assesses severity profiles in the following domains: medical, employment, alcohol, drug, psychological, legal, and family and social [34]. It has been shown to have adequate to good internal consistency, good test-retest reliability, independence across the domain composite scores, and agreement with the longer version of the ASI-Lite [34].</td>
</tr>
<tr>
<td>OAT continuation</td>
<td>Randomized clinical trial</td>
<td>OAT continuation will be assessed via a single question (“Are you currently receiving OAT (Yes or No)? Please explain”).</td>
</tr>
<tr>
<td>Length of newborn hospital stay and readmission</td>
<td>Randomized clinical trial</td>
<td>These will be determined by single questions: length of newborn hospital stay (“how many days did your newborn stay in the hospital?”) and newborn hospital readmission (“has your newborn been readmitted to the hospital for any reason after discharge? If yes, how many times? Please list reasons for each readmission”).</td>
</tr>
<tr>
<td>Perinatal maternal-fetal attachment</td>
<td>Randomized clinical trial</td>
<td>Prenatal maternal-fetal attachment will be measured with the Maternal Antenatal Attachment Scale [35,36], a reliable and validated 19-item measure of maternal-fetal attachment that includes several response formats and assesses 2 dimensions: 1) quality of attachment (11 items) and 2) intensity of preoccupation (8 items). The total scores range from 19 to 95, with higher scores indicating a higher level of attachment to the fetus [35-39]. Maternal-newborn bonding will be measured via MPAS [40,41]. It consists of 19 items assessing 3 dimensions: pleasure in interaction with the infant (5 items), absence of hostility toward the infant (5 items), and quality of mother-infant attachment (9 items). Response categories range from 2- to 5-point scales, for different items. The total score ranges from 19 to 95, with higher scores indicating higher maternal postpartum attachment to the baby. The MPAS has been found to have an acceptable level of reliability (Cronbach’s α ranging from .75 to .79; test-retest reliability r=0.86; P&lt;.001) [37,40,41].</td>
</tr>
<tr>
<td>Maternal birth experience</td>
<td>Randomized clinical trial</td>
<td>The Birth Satisfaction Scale-Revised, a 10-item, Likert-type birth satisfaction questionnaire that measures experiences of childbearing, stress, quality of care, and women’s attributes, was psychometrically validated in the United States by our research team [42-44]. Its response categories range from 1=strongly disagree to 5=strongly agree, with higher scores indicating higher birth satisfaction.</td>
</tr>
</tbody>
</table>
Study stage and outcome | Study design | Description of measures
--- | --- | ---
Maternal Depression | Randomized clinical trial | Assess via the PHQ-9, a psychometrically validated 9-item measure used to assess depression in a variety of populations [45-47]. The PHQ-9 asks participants to report on the degree they were bothered by 9 symptoms over the past 2 weeks (ie, “little interest or pleasure in doing things” and “feeling down, depressed, or hopeless”), with response categories ranging from 0=not at all to 3=nearly every day and higher scores indicating greater levels of depression symptomology.

Maternal stress | Randomized clinical trial | PSi measures parenteral stress associated with the perception of having a difficult child or a dysfunctional parent-child relationship and consists of 36 items that are rated on a 5-point Likert scale (from 1=strongly agree to 5=strongly disagree), with higher scores indicative of less total stress [48]. PSi has been shown to possess good psychometric properties and has been validated in numerous samples, including high-risk families [49]. It includes parental distress, parent-child dysfunctional interaction, and difficult child subscales, all of which will be considered individually. The child and parent domains can and will be combined to form a total stress scale score.

Breastfeeding | Randomized clinical trial | Assessed via several questions “Are you currently breastfeeding? If yes, ‘How often do you breastfeed your baby?’; if no, ‘How long did you breastfeed your baby?’”

Frequency of NAS tool use | Randomized clinical trial | Throughout their third trimester and postpartum, we will also send participants a brief weekly web-based survey link asking about the previous week’s frequency of use of the adapted mobile NAS tool and which specific modules participants viewed most, if any, to track weekly frequency of use and preference of modules.

Inclusion Criteria
Participants will be pregnant women (1) who are in the third trimester currently in OAT treatment for OUD, (2) aged >18 years, and (3) who are able to speak and understand English. Participants who did not meet the eligibility criteria will be excluded.

Data Analysis
Descriptive analyses will be conducted using percentages, means, and SDs to describe the feasibility and acceptability of the adapted mobile NAS tool for pregnant women receiving OAT. All descriptive analyses will be conducted using STATA version 14.2. This developmental stage will not require limited inferential tests because the sample sizes are small, and the qualitative work will be the primary driver of the modifications to be made and eventually tested in a randomized trial in stage 3.

Stage 3: Clinical Trial to Test the NAS Instructional Tool Versus Usual Care to Potentially Improve Postpartum Maternal-Newborn Health Outcomes

Data Collection and Inclusion Criteria
Similar to previous stages, we will recruit 30 pregnant women in their third trimester from the same 2 treatment centers to participate in stage 3 of the study. After a brief phone screening, eligible participants will then be consented and randomized 1:1 into either the intervention condition (ie, adapted NAS tool intervention, as given in the Study Intervention section) or the control condition (ie, treatment as usual [TAU], as given in the Study Intervention section). Participants will be asked to provide demographic and baseline information at this point. Women in the intervention condition will go through the mobile-based NAS instructional tool at least once during pregnancy, with their choice of going through the modules gradually while waiting at the OAT clinic to receive their dose or by scheduling a time to review the modules with research staff outside of the clinic (either at university offices or via a video conference). Participants will be met with a research assistant who will provide them with an iPad and access to the modules (via a private YouTube link), and the research assistant will be available for any questions throughout the review. Participants will be able to create a YouTube account on the iPad and will be able to access, skip, pause, and continue the modules at any time during the duration of the study through the 12-week follow-up period (ie, access after giving birth as well). They will also be able to post questions or comments under private YouTube videos for research staff to see and address. Each participant will also be scheduled for the 3 follow-up appointments (at 4, 8, and 12 weeks postpartum). Follow-up appointments will consist of filling out surveys (Table 1). Participants will be able to complete the survey on their phones or iPads. For completing all study appointments, participants will receive a total of US $125 in Amazon e-gift cards (US $25 after providing demographic and baseline information, US $25 after reviewing the modules with a research assistant, and an
additional US $25 Amazon gift card for each of the 3 follow-ups). The inclusion criteria are as follows: (1) a pregnant woman currently undergoing OAT treatment for OUD, (2) age >18 years, and (3) ability to speak and understand English. Exclusion criteria are as follows: recurring (eg, daily or almost daily) thoughts of harming themselves or others in the past 2 weeks.

**Study Intervention**

TAU pregnant women in this condition will receive care as usual, which involves continued enrollment in OAT and continued obstetric care. We will also provide them with a printed handout containing information on NAS and local resources. This level of information meets or exceeds what most mothers in this situation usually receive. Participants in the TAU condition will not receive iPads with accompanying modules; however, the handout constitutes more information than they normally receive.

**Adapted NAS Tool Intervention**

Pregnant women in this condition will receive the adapted mobile-based NAS instructional tool and TAU. Women in this condition will go through the NAS instructional tool at least once during pregnancy, with their choice of going through the modules gradually while waiting at the OAT clinic to receive their dose or by scheduling a time to review the modules. Participants will have free web access to the tool throughout their third trimester and through 12 weeks postpartum so they can access the modules at any time and as many times as desired, including after giving birth. Women will be randomized 1:1 to the intervention or TAU conditions.

**Data Collection and Outcomes**

Although the primary outcomes of focus will be maternal drug relapse (assessed via Addiction Severity Index-Lite) and OAT continuation (“Are you currently receiving OAT (Yes or No)? Please explain.”), because of the exploratory nature of this trial, several other outcomes of interest will also be assessed. At the 3 follow-up appointments, participants will complete several measures assessing numerous maternal-newborn outcomes: the 9-item Patient Health Questionnaire, a psychometrically validated 9-item measure used to assess depression in a variety of populations [45-47]; the Birth Satisfaction Scale-Revised, a 10-item, Likert-type, birth satisfaction questionnaire that measures experiences of childbearing, stress, quality of care, and women’s attributes and that was psychometrically validated in the United States by our research team [42-44]; the Parenting Stress Index short form, which measures parental stress associated with the perception of having a difficult child or a dysfunctional parent-child relationship and consists of 36 items [48]; and the Maternal Postpartum Attachment Scale (MPAS) [40,41], which measures maternal-newborn bonding. The MPAS consists of 19 items assessing 3 dimensions: pleasure in interaction with the infant (5 items), absence of hostility toward the infant (5 items), and quality of mother-infant attachment (9 items), with higher scores indicating higher maternal postpartum attachment to the baby [37,40,41]. We will also collect other outcome measures, including the length of newborn hospital stay (“How many days did your newborn stay in the hospital”), newborn hospital readmission (“Has your newborn been readmitted to the hospital for any reason after discharge? If yes, how many times? Please list reasons for each readmission.”), and breastfeeding (“Are you currently breastfeeding? If yes, ‘How often do you breastfeed your baby?’ if no, ‘How long did you breastfeed your baby?’”). Acceptability and satisfaction of the NAS tool will also be examined at follow-up by the CSQ-8 to rate overall satisfaction with the adapted NAS tool [29]. Finally, participants will also complete the 10-item SUS [30,31] (Table 2). Throughout their third trimester and postpartum, we will also send participants a brief weekly web survey link asking about the previous week’s frequency of use of the adapted mobile NAS tool and which specific modules participants viewed the most, if any, to track weekly frequency of use and preference of modules.

**Data Analysis**

Means (SDs) will be calculated for continuous variables, and percentages will be calculated for categorical variables for the 2 groups (intervention group and TAU group) at each assessment. Demographic and baseline characteristics will be tested across intervention versus TAU via independent-samples t tests and one-way analysis of variance (for continuous variables) and chi-square tests (for categorical variables). Generalized estimating equations will be used to analyze the primary longitudinal outcomes of maternal drug relapse and OAT continuation and secondary outcomes and to control for any baseline demographic differences across groups, where intervention versus TAU remains as the primary independent variable. Analyses will control for baseline and demographic outcomes and time. All inferential results will be presented as odds ratios with 95% CIs for binary outcomes and unstandardized regression coefficients with 95% CIs for continuous outcomes. We will use an α error rate of .05 as the threshold for statistical significance. All analyses will be conducted using STATA version 14.2.

**Results**

This project was funded in July 2020 (see Multimedia Appendix 1 for grant review summary statement) and approved by the institutional review board in April 2020. Data collection for stage 1 began in December 2020, and as of January 2021, we completed 18 semistructured interviews (10 with NAS providers and 8 with perinatal women receiving OAT). Common themes from all interviews will be analyzed in spring 2021 to inform the adaptation of the NAS caregiving tool. Results from stage 1 are expected to be published in summer 2021. Stage 2 data collection will commence in fall 2021, followed by the randomized controlled trial in stage 3 in late 2022.

**Discussion**

**Conclusions**

This study is highly innovative in several ways. First, despite the need to educate pregnant women at risk for delivering NAS-affected newborns and equip them with the skills necessary to successfully care for a newborn with NAS when they are pregnant, no published evidence-based interventions exist to
prepare future mothers of potential NAS-affected babies. This study will be the first to adapt an existing mobile NAS tool and develop an intervention for use with high-risk pregnant women and may be the first NAS-related intervention strategy designed to be administered to pregnant women in an OAT setting. Second, if effective, the mobile nature of our new tool will be readily scalable and well suited to tailoring for other populations of high-risk women (ie, with alcohol and or SUDs) of reproductive age. Moreover, such a tool will also be easy to update and modify, as more evidence emerges for how best to treat NAS. Although the tool is still too new to evaluate its cost efficiency, this study will prepare this tool for future analysis in the context of a larger study.

**Strengths and Limitations**

The scope of the project is unavoidably limited. First, the relatively small number of sites and pregnant women participating in stage 3 hinders generalizability and the ability to determine effectiveness. However, this novel and critically relevant trial capitalizes on existing investments to allow us to gather data on efficacy and inform a fully powered randomized controlled trial. Second, recruitment below expectation and attrition are always possible. If recruitment becomes difficult, we will hold problem-solving meetings with program staff and research team on recruitment and study advertisement strategies. However, given our proposed study sample of n=30, we foresee no issue with obtaining our target enrollment across the 2 recruitment sites. Third, although attrition can be high in high-risk populations, our participants will already be embedded within the OAT system and, therefore, already engaged with services. All activities outlined in the proposed application to develop and evaluate the mobile NAS caregiving intervention are accompanied with enhancing and evaluating the contextual fit (eg, acceptability and appropriateness) of the NAS caregiving intervention to support the application and scalability of this tool. This new tool will be designed to remain flexible to novel scientific breakthroughs in this domain such that new and modified modules can be easily created and integrated into this unique educational platform.

**Acknowledgments**

This work was supported by funding from the National Institute on Drug Abuse (K01DA051780; PI: EB). This funding source had no role other than financial support.

**Authors’ Contributions**

All authors materially participated in the research and manuscript preparation and contributed significantly to the manuscript. In addition, all authors have read and approved the final manuscript. EB developed the concept and design of the study and drafted the first version of the study protocol and manuscript. SM, HJ, JR, CBL, and RJ contributed to the study design and revised the subsequent versions of the study protocol and manuscript. OB helped with institutional review board documentation for the study protocol, data collection, and manuscript draft preparation and revision. All authors have read and approved the final manuscript.

**Conflicts of Interest**

SM and JR received research funding from the Bristol-Myers Squibb Foundation. SM received research funding from Ringful Health, LLC. SM also received research funding from Orthopedic Specialty Institute and consulted for the Consistent Care company. This funding was not related to the investigation reported here. None of the other authors have any financial, personal, or other type of relationship that would cause a conflict of interest that would inappropriately impact or influence the research and interpretation of the findings.

**Multimedia Appendix 1**

Grant review summary statement from the NIH.

[PDF File (Adobe PDF File), 145 KB - resprot_v10i4e27382_app1.pdf ]

**References**


https://www.researchprotocols.org/2021/4/e27382
45. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]

Abbreviations

CSQ-8: 8-item Client Satisfaction Questionnaire
ESC: Eat, Sleep, and Console
MPAS: Maternal Postpartum Attachment Scale
NAS: neonatal abstinence syndrome
NICU: neonatal intensive care unit

https://www.researchprotocols.org/2021/4/e27382
**OAT:** opioid agonist therapy  
**OUD:** opioid use disorder  
**SUD:** substance use disorder  
**SUS:** Systems Usability Scale  
**TAU:** treatment as usual

Edited by T Derrick; submitted 05.03.21; this is a non–peer-reviewed article; accepted 08.03.21; published 15.04.21.

Please cite as:
Burduli E, Jones HE, Brooks O, Barbosa-Leiker C, Johnson RK, Roll J, McPherson SM  
Development and Implementation of a Mobile Tool for High-Risk Pregnant Women to Deliver Effective Caregiving for Neonatal Abstinence Syndrome: Protocol for a Mixed Methods Study  
JMIR Res Protoc 2021;10(4):e27382  
URL: https://www.researchprotocols.org/2021/4/e27382  
doi:10.2196/27382  
PMID:33856360

©Ekaterina Burduli, Hendrée E Jones, Olivia Brooks, Celestina Barbosa-Leiker, Ron Kim Johnson, John Roll, Sterling Marshall McPherson. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 15.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
A Gamified, Social Media–Inspired, Web-Based Personalized Normative Feedback Alcohol Intervention for Lesbian, Bisexual, and Queer-Identified Women: Protocol for a Hybrid Trial

Sarah C Boyle¹, PhD; Joseph W LaBrie¹, PhD
HeadsUp Labs, Department of Psychology, Loyola Marymount University, Los Angeles, CA, United States

Corresponding Author:
Sarah C Boyle, PhD
HeadsUp Labs
Department of Psychology
Loyola Marymount University
1 LMU Drive
Suite 3787
Los Angeles, CA, 90045
United States
Phone: 1 8182909940
Email: sarah.boyle@lmu.edu

Abstract

Background: Sexual minority women are more likely to drink alcohol, engage in heavy drinking, and experience alcohol-related problems than heterosexual women. However, culturally tailored interventions for this population have been slow to emerge.

Objective: This type 1 effectiveness-implementation trial examines the feasibility and efficacy of a gamified, culturally tailored, personalized normative feedback (PNF) alcohol intervention for sexual minority women who psychologically identify as lesbian, bisexual, or queer (LBQ).

Methods: The core components of a PNF intervention were delivered within LezParlay, a fun, social media–inspired, digital competition designed to challenge negative stereotypes about LBQ women and increase visibility. The competition was advertised on the web through social media platforms and collaboration with LBQ community organizations. After 2 rounds of play by a large cohort of LBQ women, a subsample of 500 drinkers already taking part in the competition were invited to participate in the evaluation study. Study participants were randomized to receive 1 of 3 unique sequences of PNF (ie, alcohol and stigma coping, alcohol and control, or control topics only) over 2 intervention rounds. Randomization was fully automated by the web app, and both researchers and participants were blinded.

Results: Analyses will evaluate whether PNF on alcohol use reduces participants’ drinking and negative consequences at 2 and 4 months postintervention; examine whether providing PNF on stigma-coping behaviors, in addition to alcohol use, further reduces alcohol use and consequences beyond PNF on alcohol alone; identify mediators and moderators of intervention efficacy; and examine broader LezParlay app engagement, acceptability, and perceived benefits.

Conclusions: This incognito intervention approach is uniquely oriented toward engaging and preventing alcohol-related risks among community populations of LBQ women who may view their heavy drinking as normative and not in need of change because of the visibility of alcohol use in sexual minority community spaces. Thus, this intervention strategy diverges from, and is intended to complement, more intensive programs being developed to meet the needs of LBQ women already motivated to reduce their consumption.

Trial Registration: ClinicalTrials.gov NCT03884478; https://clinicaltrials.gov/ct2/show/NCT03884478
International Registered Report Identifier (IRRID): DERR1-10.2196/24647

(JMIR Res Protoc 2021;10(4):e24647) doi:10.2196/24647

KEYWORDS
sexual minority women; alcohol; intervention; social norms; gamification; protocol; mobile phone
Introduction

Background

The Institute of Medicine first identified sexual minority women (SMW) as a medically underserved population that disproportionately engage in hazardous drinking and carry the burden of alcohol dependence two decades ago [1]. At present, sexual orientation–based alcohol-use disparities remain with national survey data, revealing lesbian and bisexual women to be among the heaviest drinking female populations [2-4]. For instance, compared with heterosexual women of the same age, lesbian and bisexual women are more than twice as likely to engage in heavy drinking [4,5], 4-7 times more likely to meet the Diagnostic and Statistical Manual of Mental Disorders criteria for alcohol dependence [5,6], and 9-11 times more likely to report that their drinking has led to serious interpersonal consequences [5,6]. In addition, relative to their sexual majority female peers, SMW are more likely to continue patterns of heavy drinking as they age [7,8], increasing their risk for several cancers [8,9] and cardiovascular diseases [10,11]. Despite these striking disparities and consequences, targeted alcohol intervention and prevention efforts for SMW have been slow to emerge [12,13].

Recent research suggests that personalized normative feedback (PNF) may represent a welcomed low-risk, high-reward strategy for reducing alcohol use among SMW, particularly those who psychologically identify as lesbian, bisexual, or queer (LBQ) [14-16]. Founded in the social norms approach [17,18], modern PNF interventions first prompt nontreatment-seeking members of a target group or community to answer a series of web-based survey questions about their perceptions of peers’ alcohol use and their own drinking. Group members then receive individualized digital reports that use bar charts to highlight discrepancies among their perceptions of peers’ drinking, peers’ actual drinking, and their own drinking [19]. To increase the appeal and cost-effectiveness of PNF for this population, we introduced a culturally tailored, social media–inspired, gamified intervention format designed to simultaneously engage heavy drinking LBQ women just as well as their lighter-drinking and alcohol-abstaining peers and remedy the oft-cited attentional and motivational limitations associated with traditional web-based PNF intervention formats.

Why a Brief, Social Norms–Based Alcohol Intervention for SMW?

Although social norms are among the most predictive and commonly targeted antecedents to alcohol use in other heavy drinking populations [20-23] and a growing body of the literature identifies sexual minority–specific peer substance use norms as appropriate targets for intervention and prevention efforts [14,15,24-27], the dominant perspective for understanding heavy drinking among SMW is not that of social norms but rather of sexual minority stress [28-30]. This model emphasizes the role of stigma in SMW’s increased alcohol use, explaining that the increased drinking and dependence among SMW may derive from separate and combined effects of distal stressors, including experiences of prejudice, rejection, harassment, discrimination, and violence originating from a heterosexual society [28,30,31], as well as proximal stressors rooted within the individual, including internalization of stigma, concealment, and inadequate or problematic coping. Growing support for the link between sexual minority stress exposure and coping-motivated drinking [32-38] has encouraged the development of intensive, sexual identity–affirming programs that seek to reduce alcohol use and improve mental health by increasing individuals’ understanding of stigma-related processes and bolstering their adaptive coping skills and resources [39,40]. Despite the promise of these interventions for SMW seeking treatment for their problematic drinking [41,42], these strategies are unlikely to engage the larger population of SMW who do not view their drinking as problematic or lack motivation to reduce their consumption.

In contrast, web-based PNF interventions have the potential to cost-effectively reach and motivate reductions in drinking among SMW not seeking treatment and those who do not yet view their drinking as excessive or a risk to their overall health. Similar to the heavy drinking populations of college students [20,43], military personnel [21,22], and working adults [44] commonly targeted by web-based PNF alcohol interventions, bisexual- and lesbian-identified SMW have been found to overestimate descriptive peer drinking norms [13,14]. Over time, these perceptions of norms have been found to relate to drinking behavior among lesbian and bisexual SMW in the standard reciprocal feed-forward fashion observed in other groups for whom PNF has been effective [15]. Building on these findings, the primary aim of this trial is to examine the extent to which PNF designed to correct sexual identity and age-specific drinking norms is efficacious in augmenting these normative perceptions, thereby reducing LBQ women’s alcohol use and negative consequences. Recent research also suggests that in addition to general peer drinking norms (ie, frequency or quantity of consumption), this population also overestimates peer norms specific to coping-motivated drinking following collectively experienced sexual minority stressors such as the Pulse Nightclub shooting [45] and the 2016 US presidential election [46], with these misperceptions also contributing to their current and future drinking beyond the self-reported stress impact of these events. As such, the secondary aim of this trial is to evaluate whether delivering PNF on stigma-coping behaviors, in addition to alcohol use, further reduces alcohol use and consequences beyond PNF on alcohol use alone.

Testing the Boundaries of PNF Efficacy Among SMW

Few longitudinal studies have simultaneously examined the perceptions of sexual identity–specific substance use norms and concurrently experienced sexual minority stressors as predictors of substance use [47], and no published studies investigating alcohol use among SMW have considered how these predictors may influence or interact with one another over time. Rather, sexual minority stress and stigma have long been positioned in the literature as primary targets for substance use interventions, and the implicit perspective among many lesbian, gay, and bisexual health researchers seems to be that social disadvantage and stigma-related processes may render evidence-based intervention strategies that are effective in other populations ineffective among sexual minorities. Indeed, as the sexual minority stress model [30] positions sexual minority
communities as stress-buffering resources and does not consider how perceptions of substance use norms may be artificially inflated by these social environments, it is certainly possible that sexual minority stress and stigma-related processes may render a PNF alcohol intervention ineffective.

However, 2 longstanding social psychological theories support the prediction that correcting misperceived sexual identity–specific drinking norms may be maximally effective among SMW who experience severe interpersonal sexual minority stressors such as violence and harassment because of the sexual minority status. Specifically, self-categorization theory [48-50] contends that when an intergroup threat is experienced in a self-relevant domain, as would be the case when an LBQ woman experiences prejudice, harassment, or violence because of her sexual minority status, she will be particularly likely to turn to perceived in-group norms to guide her behavior. Similarly, terror management theory [51-53] posits that embracing the cultural standards and norms of a self-relevant group can protect against the deeply rooted fears of mortality likely to arise from such threatening intergroup experiences. Thus, following from both theories, to the extent that an LBQ woman experiences prejudice or victimization because of her sexual identity and has inflated perceptions of how much same-identity peers drink, she may be especially likely to increase her drinking as a means of conforming to this inflated normative standard. However, if PNF were to correct her misperceptions of these norms, she should also be especially motivated to align her drinking to the risk-reducing true norm. This study directly examines whether interpersonal stigma exposure moderates the efficacy of PNF designed to correct misperceived norms for LBQ peers’ drinking and coping behaviors.

Additional potential moderators of PNF intervention efficacy examined in this trial include preintervention alcohol consumption and negative consequences, sexual identity, age, race or ethnicity, and relationship status. Relative to lighter drinkers, heavier drinkers naturally have greater room for behavior change following PNF’s correction of drinking norms [19,22]. Thus, heavier drinking SMW are expected to exhibit larger reductions in their drinking post-PNF relative to their lighter-drinking peers. Although sexual identity, race or ethnicity, age, and relationship status are factors associated with variability in alcohol consumption among SMW [54], to date, no published studies have compared the relative sizes of discrepancies between the perceived and actual norms or the strengths of relationships between the perceived norms and alcohol consumption among subgroups of SMW who differ in these demographic characteristics. As such, this trial also evaluates whether these demographic characteristics moderate PNF intervention efficacy among LBQ women.

Addressing PNF Intervention Limitations and Engaging a Hard-to-Reach Population

An extensive study of web-based PNF interventions among college students suggests that these interventions lead to reliable but relatively short-term and modest reductions in drinking [20]. Researchers have identified several issues that, if remedied, could considerably increase the impact of this strategy. In particular, doubts about the credibility of actual drinking norms derived from previously collected data sources [55,56], defensive reactions among heavy drinkers [57,58], general inattention to feedback [59,60], and low motivation among participants [61] have been proposed as barriers to greater public health impact. The real-world suitability of this approach has also drawn criticism [62], as researchers have struggled to implement web-based PNF interventions as well as engage and retain heavy drinkers outside of study settings where participation is mandatory or participants are offered compensation at the point of recruitment [61,63]. Beyond these issues, sexual minority communities present unique challenges for intervention dissemination. Unlike universities and military bases where mandatory PNF interventions can easily target new cohorts of students and recruits, there is no single institution from which SMW can be easily recruited. In contrast, SMW must hear about community programs and judge whether the programs are credible and worthwhile.

One tool commonly used to increase the credibility and appeal of health promotion programs for minority populations is cultural tailoring, which refers to the development of interventions, messages, and materials to conform with specific cultural characteristics of the target group [64]. Recommended cultural tailoring practices for SMW include the development of programs and materials that reflect the social identities, values, and lived experiences of LBQ women as well as the involvement of LBQ community members and trusted community organizations in program promotion and delivery [12,65,66]. Following these recommendations and seeking to bolster intervention relevance, engagement, and motivation, PNF designed to correct drinking and stigma-copying norms were delivered within a larger digital competition called LezParlay. This competition was strategically crafted to reflect deep-structure cultural themes, including community members’ awareness of negative LBQ stereotypes [67-69], desire for increased identity visibility [70-72], and enjoyment of intracommunity competition and sport [73-75]. Consistent with the recommendations for surface-structure intervention tailoring [12,65,76], the LezParlay competition was also developed by an LBQ woman in the target age range and jointly promoted on the web by 4 collaborating community organizations (ie, HER social app, Autostraddle, Lez Do Brunch, and the Los Angeles Lesbian, Gay, Bisexual, and Transgender [LGBT] Center) trusted as sources of health and social information by LBQ women. In addition to cultural tailoring, LezParlay draws upon the self-determination theory (SDT [77,78]) and the nascent gamification literature [79-82] to leverage 4 evidence-based game mechanics (ie, copresence, a system of points, user-generated content, and chance-based uncertainty) to both remedy the limitations associated with traditional PNF intervention formats and foster basic psychological needs for relatedness, competence, and autonomy in this population (see Multimedia Appendix 1 [83-102] for an overview of LezParlay game mechanics and supporting literature).

One round of LezParlay was played monthly over an 8-month period, with a variable cash prize awarded monthly to the top scoring player exhibiting the greatest accuracy in their perceptions of LBQ peers. During the first 3 weeks of each
round, players were invited to size-up fellow LBQ players by browsing their social media–like profiles, submit guesses about negative stereotype-related behaviors and experiences of age- and sexual identity–matched LBQ peers (eg, What percentage of [lesbian/bisexual/queer] players in their [20s/30s/40s/50s+] own a pair of Birkenstocks? How many days per week does the typical [lesbian/bisexual/queer] player in her [20s/30s/40s/50s+] drink?), select an amount of points to wager on these guesses being true of other age- and sexual identity–matched players, and earn points for reporting on their own corresponding behaviors and experiences. In the last week of each month, all players received individualized, detailed results (ie, PNF) for a subset of the round’s questions. Animated charts and text detailed the accuracy of the player’s perceptions, how their behaviors and experiences compared with LBQ peers, summarized the stereotypes challenged, and provided their perceptual accuracy–based rank and score. Importantly, all actual norms featured in the detailed results were organically derived from the round-specific reports of players’ behaviors and experiences. Multimedia Appendix 1 provides additional literature supporting this innovative approach to the PNF intervention as well as detailed descriptions of the LezParlay round play and detailed results (ie, PNF screens).

This Study
This registered clinical trial sought to evaluate whether LezParlay-delivered PNF on alcohol use reduces alcohol consumption and negative consequences relative to PNF on control topics (aim 1); examine whether providing PNF on coping behaviors, in addition to alcohol use, further reduces alcohol use and consequences beyond alcohol PNF alone (aim 2); identify mediators (ie, perceived norms) and moderators (ie, interpersonal stigma exposure, baseline drinking, sexual identity, age, relationship status, and race or ethnicity) of intervention efficacy (aim 3); and examine broader LezParlay competition engagement, acceptability, and perceived benefits (aim 4). The following sections provide an overview of the trial design, LezParlay competition promotion efforts, randomized controlled trial (RCT) subsample recruitment, measures, and analysis plan.

Methods

Trial Design
Following recent recommendations for testing the real-world feasibility and impact of normative feedback interventions [62,103], LezParlay was examined through a type 1 hybrid-effectiveness-implementation trial [104,105]. That is, in contrast to recruiting LBQ women into a transparent, incentivized, alcohol intervention study, LezParlay was advertised as it would be in the real world—as a free, web-based competition designed to test LBQ stereotypes and increase visibility. Only after several rounds of play were a subsample of 500 drinkers, already participating in the competition, invited to participate in an incentivized evaluation study. These players were covertly randomized to receive 1 of 3 unique sequences of feedback (ie, alcohol+coping, alcohol only, or control only) over 2 consecutive rounds of play. Short-term reductions in norms and drinking were assessed 2 months later organically within the competition through a replay bonus, which invited players to boost their scores by guessing, betting, and reporting on alcohol use and control topics a second time. Following the competition, 4 months postintervention, evaluation study participants then completed a feedback survey assessing competition acceptability, perceived benefits, and feature requests for the next version of the competition. At the end of this survey, participants reported their alcohol use one final time.

Application Technology
LezParlay is a low-cost, device-responsive, HTML5 progressive web application integrated with Facebook Connect and Construct 3 game engine as well as a text message application programming interface, and email server. Although it provided a native app–like feel on Android and Apple smartphones, the app could be accessed on any internet-connected electronic device and did not require players to visit an app store or download any software. Instead, SMW simply accessed the LezParlay web app by URL [106] and were provided instructions for saving the web app to their computer’s desktop or smartphone’s home screen for easy access.

Competition Promotion
The competition was open to all LBQ women aged 21 years or above, regardless of birth sex. Players learned about LezParlay through 1 of 4 promotion strategies taking place over a 3-month period. First, before the launch of the first round, local SMW were invited to sign up through flyers and promotional items distributed at LBQ community events in Los Angeles (ie, a weekend brunch for queer women organized by a community group Lez Do Brunch and a queer casino night jointly organized by the Los Angeles LGBT Center and the Los Angeles Women’s Network). Next, as the first round began, marketing campaigns on the HER social app, the leading dating or social app for LBQ women, invited users in their 3 largest markets (ie, Los Angeles, New York City, and Chicago) to LezParlay via push notifications and in-app advertisements. An advertisement was also placed in the electronic newsletter of Autostraddle, the leading independently owned news website for queer women. During the first 3 rounds of the competition, targeted campaigns on Facebook and Instagram (Facebook Inc) also advertised LezParlay to LBQ women residing in the United States. All recruitment materials were linked to LezParlay’s informational landing page [107]. This page presented an overview of the competition and provided a sign-up button that redirected interested women to view and accept the terms of service and privacy policy (basic consent for competition participation) before creating an account. The institutional review board of Loyola Marymount University approved all recruitment materials, procedures, and intervention materials (protocol number LMUIRB2018SU14).

Procedure
After consenting to take part in the competition, users were prompted to link a valid mobile phone number to their account, and they could elect to log in with a unique email address and password combination or use their existing Facebook credentials. Next, users created their LezParlay public profile, which included a username of their choice and their sexual
identity, age group, relationship status, and pronouns. Users also had the option of uploading a profile photo or Bitmoji to represent them; entering a brief textual self-description; and connecting their Facebook, Twitter, and/or Instagram accounts so that other players could learn about them. Following account creation, players were directed to a home screen that displayed a timer counting down to the close of the current round as well as buttons to play the current round, browse player profiles, submit and vote on the questions to be parlayed in future rounds, view round winners and leaderboards, edit public profile, and change account settings. The specifics of round play and the format of the detailed results (ie, PNF) delivered at the end of each round are detailed in Multimedia Appendix 1.

RCT Subsample Recruitment

There was no upper limit on the number of SMW who could take part in LezParlay, and new players were accepted on a rolling basis throughout the competition. We aimed to recruit a minimum of 1200 LBQ women to sign up during the first 2 monthly rounds to ensure that meaningful and stable sexual identity– and age group–specific actual norms for drinking and coping behaviors could be delivered in intervention rounds 3 and 4. From this larger pool of players, 500 drinkers were recruited to participate in a LezParlay evaluation study (RCT) during the third month of play. Acting as baseline (T1) for the RCT, round 3 featured questions about alcohol use, stigma experiences, and a group of nonhealth-related control questions submitted by players. Upon submitting answers to alcohol-related questions in round 3, players were covertly screened for evaluation study eligibility based on their answers (ie, number of drinking days per week and peak drinks on a single day during the past 2 months) as well as their geolocation and the number of previous rounds played. Those who played at least one previous round, were in the United States, and reported drinking alcohol on 3 or more days per week or having 3 or more drinks on their peak drinking occasion were invited to take part in the evaluation study at the end of the round. Interested potential participants advanced to an informed consent screen that explained that the goal of the study was to evaluate the impact and format of detailed results received in LezParlay and gather player feedback to inform the next version of the competition. The information further detailed that participation in the evaluation study simply involved playing and viewing detailed results in subsequent rounds and completing a brief feedback survey at the end of the competition. Participants could earn up to US $40 in electronic gift cards of their choice to play subsequent rounds and complete the feedback survey. Those who checked a box indicating that they understood what the study participation entailed and desired to participate were welcomed into the study as LezParlay official testers.

RCT Design, Randomization, and Debriefing

The web app’s Qualtrics integration ensured that Qualtrics Research Suite’s automated randomized, commonly used in RCTs evaluating psychosocial interventions [40,41,55], could be used to randomize evaluation study participants to a PNF condition at the point of study enrollment in round 3. Randomization determined the sequence of topics on which participants received detailed results across intervention rounds 3 and 4: alcohol+coping, alcohol+control, or control only. Members of the research team were blinded to participant condition assignment, and the study participants were not aware that any sort of randomization was taking place. Rather, when detailed results were sent at the end of each round, players were prompted to choose among 3 animated, graphical doors to determine the 1 to 2 round topics on which they would view detailed results (see Figure S2 in Multimedia Appendix 1). Although the results topics were truly determined via chance in most rounds of the competition, the doors of evaluation study participants were fixed to open to their randomly assigned feedback topics regardless of the door they selected in rounds 3 and 4. Upon completing the feedback survey at the end of the competition, participants were debriefed regarding the study’s research questions and the fixed sequences of health or control feedback they were randomized to receive in rounds 3 and 4 of the competition.

Intervention Rounds

All players taking part in the third round of LezParlay estimated the drinking behaviors of the typical same-sexual identity player in their age group during the previous 2 months, reporting on their perceptions of the typical player’s (1) maximum number of drinks consumed on a single occasion, (2) average number of drinks consumed per occasion, and (3) average number of drinking days per week [103]. Players also estimated the number of negative alcohol-related consequences experienced over the previous 2 months by a typical player in their sexual identity and age group from a list of 8 negative consequences (ie, had a hangover or illness, got in a physical or verbal fight, had problems with significant others, missed a social engagement or event, had problems with friends or family, performed poorly at work or school, had problems with money, and had an unwanted or regrettable sexual experience). Players then answered parallel items assessing their own drinking and consequences over the corresponding 2-month period.

Players taking part in round 4 of the competition were prompted to think about how other players deal with stress and sexual minority stigma and asked to estimate the percentage of time (ie, 0%-100%) a typical player in their sexual identity and age group tried to feel better during the past month by (1) drinking alcohol; (2) taking a drug; (3) meditating, using relaxation techniques, or exercising; and (4) talking to a close other or mental health professional. Players were then prompted to think about how they themselves dealt with stress and stigma and responded to parallel items. The actual norms variably delivered to evaluation study participants in PNF at the end of rounds 3 and 4 were derived by computing the actual average response of all players submitting responses in each sexual identity and age group.

Participants randomized to receive PNF on control topics received detailed results for nonhealth-related topics in round 3 (eg, household repair ability, frequency of home improvement store visits, or tool box ownership) and round 4 (eg, time in between relationships, texting exes, and partners being confused for sisters). Multimedia Appendix 1 includes a link to view an example of detailed results for a control topic and a treatment topic delivered in the intervention rounds.
RCT Measures

**Demographic and Psychosocial Covariates**

At sign up, all players reported their sexual identity, relationship status, and age group. Upon enrolling in the evaluation study, participants also reported their race, ethnicity, and actual age in years. The feedback survey at the end of the competition prompted the study participants to report their relationship status and sexual identity.

**Perceived Alcohol-Related Norms and Behaviors**

As described previously, perceived drinking norms and alcohol-use behaviors were assessed organically in competition rounds 3 (T1; baseline) and 7 (T2; 2-month follow-up) by items modeled after Baer’s quantity, frequency, and max measure [108] in combination with additional norm and behavior items, respectively, examining negative alcohol-related consequences. These items were assessed a final time at the end of the postcompetition survey (T3; 4-month follow-up). Measures at each timepoint were referenced from the previous 2-month period. As done in previous gamified PNF pilot studies with college students [43,109,110], composite measures of perceived alcohol-use norms and alcohol-use behavior at baseline and follow-up will be computed by z scoring and then averaged across respective sets of individual items at each timepoint. In addition to these composites, 3 key outcomes of interest in alcohol intervention research are to be examined individually pre- and postintervention: (1) estimated drinks per week over the previous 2 months (computed by multiplying the reported number of drinking days per week and the average number of drinks per occasion at each timepoint), (2) peak drinks on 1 occasion over the previous 2 months, and (3) the number of negative alcohol-related consequences over the previous 2 months.

**Interpersonal Stigma Exposure**

Interpersonal stigma exposure was also assessed at baseline (round 3) and follow-up (as a replay bonus topic in round 7). Players guessed about the stigma experiences of other players and reported on their own stigma experiences over the previous 2 months. Stigma-related norms were not corrected in the competition, and these perceptions were only assessed to make sense for players to report on their own recent interpersonal stigma exposure (a theorized moderator of conditional effects on drinking) at the same time that they were reporting on their alcohol use and negative consequences in the game. Players’ recent exposure to severe interpersonal stigma was assessed by their responses to 2 items: (1) “During the past 2 months, how many times have you been physically harmed due to your sexual identity?” and (2) “During the past 2 months, how many times have you been verbally harassed or threatened (online or in person) due to your sexual identity?” Associations between pairs of stigma items at each timepoint are expected, and we anticipate combining responses to derive severe interpersonal stigma scores.

Feasibility Measures

**Reach and Engagement**

Data from Google Analytics and the app’s back end will allow us to examine the total number of players who signed up to participate in the LezParlay competition in the absence of traditional study participation incentives; identify the promotional channels that brought them to the app; and detail the players’ demographic characteristics, states of residence, average number of log-ins, and number of rounds completed.

**Acceptability**

Feedback surveys prompted study participants to rate numerous aspects of the competition (the stereotype challenge concept, topics and questions, detailed results, leaderboards, the ability to browse player profiles, the ability to submit questions, the ability to bet points on the accuracy of guesses, text messages, and email communications from LezParlay) on Likert-type scales ranging from did not like at all (0) to liked very much (5).

**Perceived Benefits**

A single yes or no item asked participants whether they felt that participating in the LezParlay competition was psychologically beneficial. Those selecting yes in response were asked to enter text describing the perceived benefits.

**Improvements and Requested Features**

A final free-response item asked participants to share recommendations they had for improving the competition and describing the features they would like to see in the next version.

**Results**

**Overview**

This project was funded by the National Institute on Alcohol Abuse and Alcoholism in April 2018. The institutional review board approval was granted in August 2018, and the LezParlay app was developed per approved protocol between August 2018 and November 2018. The competition launched in January 2019, and the collection of all efficacy and feasibility data was completed in September 2019. A total of 2677 LBQ women participated in the LezParlay competition, with 500 LBQ drinkers recruited into the efficacy trial. Data cleaning and analysis were delayed by several months because of COVID-19–related delays and is underway as of January 2021. The results are expected to be published in summer 2021.

**Data Analytic Plan for Evaluating Intervention Efficacy**

An intent-to-treat approach will be used to examine LezParlay treatment effects at 2 and 4 months postintervention on 4 outcomes: composite alcohol use, estimated number of drinks per week, peak number of drinks on 1 occasion, and number of negative alcohol-related consequences. Preliminary analyses will examine potential biases related to attrition and missing data [111,112], inspect outcome distributions, and evaluate potential baseline differences among conditions. As the latter 3 outcomes are count variables (ie, estimated drinks per week,
peak drinks, and negative consequences), they are likely to be substantially skewed and best approximated by either Poisson or negative binomial distributions.

**Main Effects**

At 2 and 4 months after the delivery of treatment PNF, participants in both conditions receiving treatment PNF on alcohol use (ie, alcohol-coping and alcohol only) are expected to report reduced drinks per week, peak drinks, and negative consequences relative to those in the control PNF condition. Furthermore, participants in the alcohol-coping condition are expected to exhibit larger reductions in their alcohol use and negative consequences at postintervention follow-ups than participants in the alcohol-only PNF condition. Multilevel models (MLMs [113,114]) with full maximum likelihood specification will be used to test these predictions. Time will be specified as a level 1 varying predictor nested within individuals (level 2). Intercept treatment differences will represent treatment differences at baseline (eg, conditional differences in drinking at baseline), and slope differences will represent changes over time (eg, did participants in treatment conditions reduce their drinking between baseline and follow-up assessments more than control participants). The intercept includes a random effect, which will model the subject-specific heterogeneity in alcohol-related outcomes, thereby controlling for correlated data due to individuals. Main effect models will also control for covariates: age, sexual identity, race, ethnicity, relationship status, and severe interpersonal stigma exposure.

**Tests of Mediation and Moderation**

Tests of mediation will examine whether perceived drinking norms at the 2-month follow-up mediate relationships between condition and alcohol-use outcomes at the 4-month follow-up. PROCESS bootstrap tests [115,116] will be used to test the mediation. These models will control for baseline measures of potential mediating variables (ie, norms) and outcomes (ie, alcohol use and consequences). Moderation analyses will be examined within an MLM framework and will examine whether the efficacy of treatment PNF varied as a function of participants’ baseline drinking, sexual identity, exposure to severe interpersonal stigma, or other demographic characteristics. In the presence of significant interactions, exploratory moderated mediation models [115,117,118] may simultaneously estimate the conditional direct and indirect effects associated with the different levels of moderating variables.

**Power Analysis**

Informed by previous research examining the effects of web-based alcohol PNF on changes in normative perceptions and drinking in other populations (Cohen d=0.22 [20,119]), the comparably larger effect size revealed in a similar gamified PNF intervention for college students (Cohen d=0.46 [109]), power analyses using the standard 0.80 power of detecting a significant effect, P<.05, and an effect size of Cohen d=0.30 indicate a sample size of 375 (125 participants in each condition) to be sufficient to detect small-to-medium effects using repeated measures MLMs (ie, 2 levels, 3 arms, and randomization at the individual level) as well as tests of mediation and moderation. Thus, our sample size of 500 will allow us to detect modest effects with even 30% attrition.

**Data Analytic Plan for Evaluating Feasibility**

**Discussion**

**Trial Overview**

Joining the growing number of gamified health interventions being developed for sexual minority men and youth [121-126] and extending gamified PNF pilot work with college students [43,83,109,110], this project leverages evidence-based digital game mechanics informed by SDT to deliver PNF within a novel digital competition designed to challenge LBQ stereotypes and increase visibility. Notably, several deep-structure themes well documented among LBQ women [63-66], and social norms theory predicts that players are likely to overestimate undesirable stereotypical behaviors among peers of same-sexual identity (eg, promiscuity and infidelity among bisexual women, unhealthy relationship behaviors, and transphobic attitudes among lesbians). As such, we anticipate that revealing and reinforcing true norms for such experiences and attitudes may carry psychological benefits (ie, reducing identity-related stigma and increasing collective self-esteem) for participants beyond reduced alcohol consumption.

This hybrid trial also follows recent recommendations for improved design and evaluation of social norms–based health interventions [62], as both mediators and moderators of intervention effectiveness are examined, and, importantly, the LezParlay competition is framed and advertised as it would be outside of the study setting. Only later is an incentivized clinical trial subsample of drinkers recruited from the larger population of players already engaging with the competition (notably in the absence of traditional study participation incentives). This hybrid design allows the research team to cost-effectively assess the feasibility of drawing large numbers of SMW to the broader LezParlay competition via targeted promotional channels, player engagement with different areas of the web app, and competition acceptability and ways in which the competition might be improved. Simultaneously, recruiting a subsample of alcohol-consuming SMW already taking part in the competition into an incognito RCT allows for the evaluation of whether PNF on alcohol use and stigma-coping behaviors meaningfully

https://www.researchprotocols.org/2021/4/e24647
reduces alcohol consumption and negative consequences relative to control PNF. In sum, this design allows critical questions about feasibility and efficacy to be jointly addressed, with minimal costs to internal or external validity.

The examination of demographics in terms of both LezParlay engagement and moderators of intervention efficacy will identify the groups of SMW most engaged and impacted by this strategy. Furthermore, as previous work with SMW has not examined potential interactions between perceptions of sexual identity–specific drinking norms and experiences with violence or harassment because of sexual minority status, examining the potential interplay between these established predictors of drinking allows this project to make a significant contribution to the larger sexual minority health literature. Similarly, as normative feedback interventions have not been widely considered as potential strategies for reducing problematic substance use or other health-risk behaviors among members of stigmatized health disparity populations, determining whether minority identity–based violence and harassment makes PNF more or less efficacious is a critical research question that may also carry intervention development implications for these populations (eg, racial and ethnic minorities and gender minorities).

### Limitations and Future Directions

Funded by the National Institutes of Health’s exploratory/developmental R21 grant mechanism, the goal of this initial trial is to evaluate the feasibility and efficacy associated with the LezParlay competition as a minimally viable product taking the form of an extremely low-cost, progressive web app designed and coded by the first author, who is a member of the target population. Thus, representing the preliminary step in a larger program of LezParlay-gamified PNF intervention research, findings from this trial will not speak to the feasibility or efficacy of delivering the intervention through a more sophisticated and polished native smartphone app that would likely be more desirable and user-friendly and thus better equipped to attract and retain LBQ women. In addition, PNF is the only intervention component featured in the initial version of LezParlay and, furthermore, only static descriptive norms for drinking and coping are corrected. Findings from this study may suggest high feasibility for the gamified approach and stereotype challenge framing (ie, large numbers of SMW are engaged by the competition and participants and report psychological benefits), but treatment PNF fails to meaningfully reduce drinkers’ consumption and negative alcohol-related consequences relative to control. In this event, the stereotype challenge concept and game mechanics may be retained; however, future versions of the LezParlay app might expand PNF to correct additional types of alcohol and coping norms (ie, injunctive, affective, and dynamic norms) and/or deliver additional intervention components such as skills training around healthy coping strategies and/or local alcohol treatment information (ie, referral to treatment).

The intervention’s real-time, organic generation of actual norms from LezParlay players’ round-specific reports of their behaviors represents both a point of innovation and a potential limitation. Foremost, this approach diverges significantly from traditional PNF interventions that derive actual drinking norms from an existing data source or a separate norms documentation study conducted before the recruitment of the intervention sample [19-22]. As PNF research suggests that the practice of using previously collected actual norms data may undermine intervention efficacy by diminishing PNF credibility and interest [55-57,110], LezParlay sought to increase the psychological proximity and relatability of the peer group by making the individuals on whom actual norms are based digitally present and visible. As a point of innovation, this format may help extend PNF interventions to sexual minorities and other hard-to-reach populations for whom existing norms data do not exist and would be difficult and/or costly to collect through a separate survey study. Furthermore, trial results from a similar gamified PNF intervention for college students revealed that the actual drinking norms similarly derived in-game among visible peers differed very little from those derived from separate survey samples of students at the same university [109]. However, as this approach has not previously been tested with adult LBQ women, it is critical to establish that the actual norms derived in real time from LezParlay rounds are sufficiently risk-reducing, stable, and approximately equivalent to those derived from comparable survey samples containing the same age and sexual identity groups [14-16,91].

Another limitation pertains to this study’s organic assessment of alcohol outcomes and some moderators within the rounds of the competition. Although this is a major strength, in that it eliminates the demand characteristics that often plague transparent alcohol intervention studies and substantially decreases trial costs, this also meant that the key constructs could only be assessed by a few items, and the language of the items could not be too formal or clinical in tone. Although Baer’s frequency, quantity, and max measure [103] fit well in this regard as a short, validated measure of alcohol use, it would have also been valuable to include longer validated survey measures to more formally assess alcohol-related outcomes and screen for alcohol use disorder. In the event that LezParlay feasibility and efficacy are demonstrated in the initial trial, we anticipate seeking additional funding for a larger trial that will include traditional survey-based baseline and follow-up assessments and test an expanded set of potential mediators and moderators.

A final limitation pertains to this study’s narrow focus on the direct effect of PNF treatment on alcohol-related outcomes. Norms for other health behaviors (ie, stigma coping, smoking, exercise, and health care utilization) were also corrected within the broader competition; however, the RCT was not designed to examine potential PNF-related changes in these behaviors. Similarly, players were expected to overestimate several undesirable stereotypical behaviors among same-sexual identity peers in nonintervention rounds of the competition (eg, promiscuity and infidelity among bisexual women, unhealthy relationship behaviors, and transphobic attitudes among lesbians). Although outside the scope of this initial trial, revealing and reinforcing true norms for these experiences and attitudes in LezParlay may carry psychological benefits for SMW partaking in the competition (ie, reducing identity-related stigma, increasing feelings of belonging, and/or collective
self-esteem). Thus, assessing pre-post competition changes in these constructs and evaluating the extent to which challenging negative stereotypes through the larger LezParlay competition might buffer stigma-related processes and thereby reducing drinking and improving other health outcomes remain to be critical next steps in the larger program of research.

Conclusions

This hybrid trial will examine the efficacy and feasibility of an innovative, culturally tailored, evidence-based alcohol intervention for LBQ moderate-to-heavy drinkers, thereby narrowing the disparity in alcohol intervention research and practice. This *incognito*, gamified intervention approach is uniquely oriented toward engaging and preventing alcohol-related risks among LBQ community members who may view their heavy drinking as normative and not in need of change because of the visibility of alcohol consumption in LBQ community spaces. Thus, this intervention strategy diverges from and is intended to complement more intensive intervention programs being developed to meet the needs of SMW already motivated to reduce their consumption and those seeking culturally tailored treatment for alcohol-use disorder and comorbid mental health problems [40,41].

Acknowledgments

Manuscript preparation was supported by developmental/exploratory grant R21AA025767-01A from the National Institute on Alcohol Abuse and Alcoholism. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute on Alcohol Abuse and Alcoholism or its affiliates.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of LezParlay game mechanics, supporting literature, round play, and detailed results.

[DOCX File, 490 KB - resprot_v10i4e24647_app1.docx]

Multimedia Appendix 2
Peer-review report by the National Institute on Alcohol Abuse and Alcoholism.

[PDF File (Adobe PDF File), 186 KB - resprot_v10i4e24647_app2.pdf]

References


46. Boyle S, Costine L, LaBrie J. Cross-sectional and forward-feedback relationships between peer coping norms and LGB adults own coping-motivated alcohol use following stressors. Alcoholism Clinical Experimental Research 2017;41:259A.


Abbreviations

- LBQ: lesbian, bisexual, or queer
- LGBT: lesbian, gay, bisexual, and transgender
- MLM: multilevel model
- PNF: personalized normative feedback
- RCT: randomized controlled trial
- SDT: self-determination theory
- SMW: sexual minority women
a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
A Technology-Based Physical Activity Intervention for Patients With Metastatic Breast Cancer (Fit2ThriveMB): Protocol for a Randomized Controlled Trial

Siobhan Phillips¹, MPH, PhD; Payton Solk¹, BA; Whitney Welch¹, PhD; Lisa Auster-Gussman¹, PhD; Marilyn Lu¹, BSc; Erin Cullather¹, BSc; Emily Torre¹, BA; Madelyn Whitaker¹, BSc; Emily Izenman¹, BA; Jennifer La¹, BSc; Jungwha Lee¹, PhD; Bonnie Spring¹, PhD; William Gradishar¹, MD

Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Siobhan Phillips, MPH, PhD
Northwestern University Feinberg School of Medicine
680 N Lake Shore Drive
Suite 1400
Chicago, IL, 60611-4407
United States
Phone: 1 13125034235
Email: smphillips@northwestern.edu

Abstract

Background: Increased incidence and life expectancy have resulted in a growing population of patients with metastatic breast cancer, and these patients experience high rates of morbidity and premature mortality. Increased physical activity (PA) is consistently associated with improved health and disease outcomes among early-stage survivors. However, there is a paucity of research on PA in patients with metastatic breast cancer, and existing PA interventions have exhibited low feasibility because of their focus on intense PA and/or requirement of on-site visits. Mobile health (mHealth)–based PA interventions may be particularly useful for patients with metastatic breast cancer because they allow for remote monitoring, which facilitates individual tailoring of PA recommendations to patients’ abilities and may minimize participant burden. However, no studies have examined mHealth PA interventions in patients with metastatic breast cancer.

Objective: We aim to address these critical research gaps by testing a highly tailored technology-based intervention to promote PA of any intensity (ie, light, moderate, or vigorous) by increasing daily steps in patients with metastatic breast cancer. The primary aim of this study is to test the feasibility and acceptability of the Fit2ThriveMB intervention. We will also examine outcome patterns suggesting the efficacy of Fit2ThriveMB on symptom burden, quality of life, and functional performance.

Methods: The Fit2ThriveMB trial is a two-arm pilot randomized controlled trial that will compare the effects of a smartphone-delivered, home-based PA intervention and an attention-control education condition on PA and quality of life in low-active female patients with metastatic breast cancer. A subsample (n=25) will also complete functional performance measures. This innovative trial will recruit 50 participants who will be randomized into the study’s intervention or control arm. The intervention will last 12 weeks. The Fit2ThriveMB intervention consists of a Fitbit, coaching calls, and the Fit2ThriveMB smartphone app that provides self-monitoring, a tailored goal-setting tool, real-time tailored feedback, app notifications, and a group message board. Assessments will occur at baseline and post intervention.

Results: The Fit2ThriveMB study is ongoing. Data collection ended in February 2021.

Conclusions: Data from this study will provide the preliminary effect sizes needed to assemble an intervention that is to be evaluated in a fully powered trial. In addition, these data will provide essential evidence to support the feasibility and acceptability of using a technology-based PA promotion intervention, a scalable strategy that could be easily integrated into care, among patients with metastatic breast cancer.

Trial Registration: ClinicalTrials.gov NCT04129346; https://clinicaltrials.gov/ct2/show/NCT04129346

International Registered Report Identifier (IRRID): DERR1-10.2196/24254

(JMIR Res Protoc 2021;10(4):e24254) doi:10.2196/24254
KEYWORDS
physical activity; metastatic breast cancer; technology; randomized controlled trial; mobile phone

Introduction

Background
Almost 155,000 women live with metastatic breast cancer in the United States as of January 1, 2017 [1]. The number of women diagnosed with metastatic breast cancer is expected to increase by 31% over the next 10 years [1]. Furthermore, treatment advances have doubled the 5-year survival rates in the last two decades by 36% and 11% of women survive for ≥10 years [1]. Thus, the number of women with metastatic breast cancer is increasing. Women with metastatic breast cancer have higher rates of physical impairment [2], greater symptom burden [3], and lower levels of fitness and strength [3,4] than early-stage survivors and noncancer controls, resulting in compromised quality of life (QoL). Interventions are needed to alleviate adverse health effects and allow patients with metastatic breast cancer to function optimally in the years they survive with advanced cancer. However, few health-enhancing interventions exist for women with metastatic breast cancer, and only 2%-5% [5] of research funds for breast cancer are spent on metastatic disease despite its high morbidity rates. Thus, research in this population is urgently needed.

Increased physical activity (PA) is consistently associated with fewer treatment-related side effects, higher QoL, increased survival, and reduced recurrence and mortality among survivors of early-stage breast cancer [6-9]. Increasing light-intensity activity and reducing sedentary time may also reduce functional decline [10] and mortality [11] and improve QoL. [11,12] and body composition [13] independent of more intense activity. However, there is a paucity of research on PA in patients with metastatic breast cancer. Two recent reviews of PA interventions in mixed samples with advanced cancer indicate that PA is safe and feasible for these populations, and PA may be associated with a variety of benefits, including improved aerobic capacity, strength, physical and psychosocial function, QoL, fatigue, sleep quality, and body composition [14,15]. In addition, a small study on patients with metastatic breast cancer found that increased PA was associated with longer survival [16]. However, only 4 randomized controlled trials (RCTs) have focused specifically on PA in patients with metastatic breast cancer [3,17-19]. Although these studies demonstrated that both supervised and home-based PA programs are safe for patients with metastatic breast cancer, findings regarding effects on PA, fitness, and patient-reported outcomes were equivocal; 2 studies found no effect on the outcomes of interest [17,20], and 2 found trends toward moderate intervention effects on fatigue and physical well-being, fitness, and the 6-minute walk test [18,19]. However, sample sizes were small. The null findings were largely because of their focus on intense activity and/or requirement of on-site visits. The one study that used a wearable tracker is a single-arm study that found high adherence (96%) to using the tracker [21]. Using mobile health (mHealth) technology for PA promotion interventions may be particularly useful for these women because it allows for remote monitoring of patients, which facilitates individual tailoring of PA programs to patients’ abilities. mHealth interventions also do not require travel to on-site, supervised activity sessions, thus reducing participant burden. However, no studies have examined the effects of an mHealth intervention on objectively measured PA in patients with metastatic breast cancer in a controlled trial.

The primary goal of this study is to address these critical research gaps by testing the feasibility and acceptability of a highly tailored 12-week mHealth intervention to promote activity of any intensity (ie, light, moderate, or vigorous) by increasing daily steps using a two-arm RCT in patients with metastatic breast cancer. Secondary goals include examining the intervention’s effects compared with a control group on symptom burden and QoL in the full sample and functional performance in a subsample (n=25).

Study Aims
The Fit2ThriveMB trial was designed to pursue 3 objectives. First, we will examine the feasibility and acceptability of Fit2ThriveMB, a 12-week mHealth PA intervention in patients with metastatic breast cancer. Second, we will examine the potential effects of Fit2ThriveMB on accelerometer-assessed PA. Finally, we will examine outcome patterns suggesting the efficacy of Fit2ThriveMB in improving symptom burden (ie, fatigue, depression, anxiety, pain, and physical function) and QoL in the full sample and functional performance in a subsample.

Methods

Overview and Study Design
FFit2ThriveMB is a 12-week mHealth pilot RCT conducted to examine the feasibility and acceptability of a 12-week technology-based PA promotion program. We will also compare the Fit2ThriveMB intervention with a healthy lifestyle attention-control group and examine outcome patterns, suggesting the efficacy of Fit2ThriveMB for increasing PA and improving symptom burden, QoL, and functional performance. We randomized 50 low-active patients with metastatic breast cancer to each condition. Participants assigned to the Fit2ThriveMB condition will receive a PA promotion intervention that incorporates behavior change principles based on social cognitive theory (SCT) [22] and includes a Fitbit, weekly coaching calls, and the Fit2ThriveMB smartphone app, which includes tools for self-monitoring and tailored goal setting, real-time tailored feedback, app notifications, and a social feed. Participants in the healthy lifestyle control will be instructed to download a commercially available cancer care app (Cancer.Net) and will be provided with educational materials and support calls. The participants will complete assessments at baseline and postintervention (12 weeks). Participants in the control condition will receive the Fit2ThriveMB app and a Fitbit following completion of the
12-week assessments. An overview of participant flow through the study is shown in Figure 1.

**Figure 1.** Fit2ThriveMB participants flow through the study.

Conceptual Model

SCT [22] is well recognized as a useful framework for designing PA interventions for cancer survivors [23,24]. SCT specifies a core set of determinants (self-efficacy, goal setting, facilitators and barriers, and outcome expectations) and the mechanisms by which they work [25]. Specifically, SCT postulates that self-efficacy is both directly and indirectly related to PA, via facilitators and barriers (ie, lack of access to facilities or social support), goal setting and self-regulation (ie, monitoring PA or using feedback to measure progress) and outcome expectations (ie, belief that PA will result in a specific outcome) [25]. Figure 2 details our conceptual model of how we hypothesize that Fit2ThriveMB will influence PA, symptom burden, QoL, and functional performance. On the basis of our conceptual model, we hypothesize that the Fit2ThriveMB intervention components will improve self-regulatory and goal-setting skills, increase self-efficacy via increased mastery experiences and social persuasion, increase realistic outcome expectations, improve facilitators by increasing social support, and increase the ability to overcome barriers, resulting in increased PA. Increased PA is, in turn, hypothesized to result in favorable changes in symptom burden, QoL, and functional performance.
Recruitment and Screening

Eligibility

The inclusion criteria are being female, aged ≥18 years, diagnosed with metastatic breast cancer (stage IV), with a life expectancy of ≥12 months, low active (ie, self-reported engagement in <150 minutes of moderate-to-vigorous physical activity [MVPA] per week), fluent in spoken and written English, owning a smartphone, and having access to a computer with internet access to complete assessments. The exclusion criteria include untreated brain metastases, uncontrolled cardiovascular disease or other major contraindications (ie, nonambulatory or severe cognitive or functional limitations) to activity participation, and current enrollment in another dietary or activity trial. All participants must obtain medical and diagnostic eligibility clearance from their oncologists. Diagnosis will also be confirmed via medical records before enrollment.

Recruitment Procedures

Women will primarily be recruited from oncology clinics at a large Midwestern academic medical center via medical records and physician referrals. Patients will also be recruited via support groups, fliers, and local cancer events. All potential participants will be sent a study recruitment email, providing a brief description of the study and assessing patient interest. If interested, a personalized link in the email will automatically redirect the potential participant to complete the web-based screening. Potential participants will be sent 2 recruitment emails. If no response is received to these contact attempts, after 1-2 weeks, the study team will follow up with a phone call. If the individual answers and is interested in participating, the study team will describe the study and screen the individual via phone or resend them the link to the web-based screener, depending on patient preference. If the potential participant does not answer the phone call, the study team will leave a voicemail and send 1 final recruitment email with a link to schedule a recruitment call.

Screening

All interested women will complete the initial eligibility screening on the web or over the phone. The screening will take approximately 15 minutes and includes an evaluation of all eligibility criteria, including self-reported PA and completion of the Physical Activity Readiness Questionnaire [26], which assesses cardiovascular disease history, symptoms, risk factors, and other health issues. All participants must obtain medical and eligibility clearance from their oncologist to participate. Diagnosis will also be confirmed via medical records before enrollment. The subsample of participants who participate in functional performance tests will also be asked to complete the Falls Risk Questionnaire [27], and the results will be shared with the oncologist to inform clearance decisions and used by the study staff to prepare for any potential balance issues that may arise during the functional performance test. Eligible candidates will be emailed a copy of the informed consent and a study overview document to review and will be scheduled for a recruitment call.

Recruitment Call

During the recruitment phone call, staff will explain the study in greater detail and confirm that participants meet the eligibility criteria for cancer diagnosis, their activity level and if they have a major contraindication to PA participation, and their intention to participate in the study. The call will last, on average, from 15 minutes to 20 minutes. If the potential participant is screened over the phone, the recruitment call will be completed as part of the screening phone call unless the individual prefers to complete it at a later date. Interested and eligible women will be emailed a link to a web-based version of the informed consent and a copy of the permission to contact the physician form. They will be instructed to complete and sign both documents.
and submit them on the web. Once both documents are complete, the study team will acquire physician consent.

**Data Collection**

All participants who consent to participate will be shipped an assessment kit. The kit will include an accelerometer, accelerometer instructions, and an accelerometer log for all participants. Participants will be instructed to wear the accelerometer for 7 consecutive days (see Physical Activity Measures for further details) and return the accelerometer to the study team via the provided postage-paid return envelope. A personalized link to a battery of questionnaires will be emailed to participants. The participants will complete the questionnaires at home and submit them on the web via RedCap before their videoconference. These procedures will be the same at both testing time points.

For the subgroup participating in the functional performance test, the assessment kit will also include a functional performance test kit with instructions for preparing for their videoconference functional performance testing at baseline (see Physical Function Measures for further details). Following baseline functional performance testing, participants will be instructed to put the functional performance test kit in a safe-keeping spot to reuse at 12 weeks. After the 12-week testing, they will be instructed to send the test kit materials back.

Participants will be regularly reminded via phone or email to wear the accelerometer, answer the web-based questionnaires, and attend their functional performance testing sessions. Participants will not be randomized until all the baseline data are complete. Participants will be incentivized for completion of assessments and allowed to keep the Fitbit (intervention) or provided with a Fitbit (control) if they complete assessments at both time points.

All Fitbit data and data on app usage will be collected throughout the duration of the intervention and will be stored in a database developed specifically for this study, which is on a secure, password-protected server only accessible to study investigators.

**Measures**

Adherence and retention will be evaluated during the intervention. All other measures will be assessed at baseline and postintervention (12 weeks).

**Feasibility and Acceptability**

The feasibility of each component will be measured during the intervention and immediately after the intervention. Measures include participant retention (number of participants who dropped out/number randomized), objective (when possible) and self-reported usage of intervention components (percentage of days of adhering to the daily goal and percentage of days of wearing Fitbit), and safety as measured by the number and severity of adverse events reported spontaneously and during nons spontaneous assessments. Fitbit is sufficiently validated for steps to support its role as an adherence measure in this study. Adherence during the 12-week intervention will be monitored continuously using the Fit2ThriveMB app. We will obtain Fitbit wear time, steps, and time spent in sedentary, and light-, moderate-, and vigorous-intensity activity. Acceptability will be measured via a process evaluation of the perceptions of experiences of patients with metastatic breast cancer with Fit2ThriveMB. This evaluation will assess the following: perceived effectiveness of intervention components, plans to continue PA and intervention use, intervention elements liked or disliked, and satisfaction with program delivery, assessments, and staff. Items will be rated on a scale ranging from 1 (worst) to 5 (best). We will also conduct semistructured postprogram interviews with individuals who participate in the intervention group to obtain feedback on each intervention component and recommendations for how the intervention can be improved.

**Physical Activity**

PA will be measured objectively using self-report measures. PA will be assessed objectively by accelerometer at baseline and 12 weeks using an ActiGraph accelerometer (model GT3X-BT, ActiGraph). ActiGraph, a valid and reliable measure of PA [28,29], has been widely used in oncology research [30,31]. Participants will be instructed to wear the activity monitor on the nondominant hip during all waking hours (except when bathing or swimming) for 7 consecutive days. The accelerometer will collect activity data in 10-second intervals (ie, epochs). Upon receipt of the accelerometer, the data will be immediately downloaded and checked for valid wear time. If there is not at least five days of valid measurement with 10 hours of valid wear time [32], participants will be asked to rewear the monitor. The average number of minutes of daily total activity and time spent in sedentary, light, moderate, and vigorous activity will be calculated using established cut-points [33,34]. Data on steps will also be collected. We are primarily interested in data on the total minutes of PA of any intensity (total weekly light, moderate, and vigorous) and steps but will also examine average daily total counts. The Godin Leisure-Time Exercise Questionnaire [35] will be used to assess self-reported minutes of PA (light, moderate, and vigorous) during each measurement period.

Fitbit will be used to collect activity data throughout the intervention, and participants will wear both Fitbit and ActiGraph during week 12 assessments. Fitbit data will not be used as an outcome measure. Instead, these data will be used to monitor adherence, as detailed below.

**Symptom Burden**

Women will complete reliable, well-validated Patient-Reported Outcomes Measurement Information System [36] self-report assessments of symptoms including fatigue, depression, anxiety, pain interference, and physical function, and the Functional Assessment of Cancer Therapy-Breast [37], a well-validated measure of QoL, via web-based questionnaires at each time point. The details of these measures are provided in Table 1.
Table 1. Study questionnaires.

<table>
<thead>
<tr>
<th>Construct and measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symptom burden [36]</strong></td>
<td></td>
</tr>
<tr>
<td>PROMIS® Physical Function Short Form 20a [36]</td>
<td>Measures functional limitations and interference over the past 7 days</td>
</tr>
<tr>
<td>PROMIS Fatigue Short Form 8b [36]</td>
<td>Measures frequency of fatigue symptoms and interference over the past 7 days</td>
</tr>
<tr>
<td>PROMIS Depression Short Form 8b [36]</td>
<td>Measures the frequency of a variety of depressive symptoms over the past 7 days</td>
</tr>
<tr>
<td>PROMIS Anxiety Short Form 8b [36]</td>
<td>Assesses self-reported fear (fearfulness and panic), anxious misery (worry and dread), hyperarousal (tension, nervousness, and restlessness), and somatic symptoms related to arousal (racing heart and dizziness) over the past 7 days</td>
</tr>
<tr>
<td>PROMIS Pain Interference Short Form 8b [36]</td>
<td>Measures the self-reported consequences of pain on relevant aspects of one’s life over the past 7 days</td>
</tr>
<tr>
<td><strong>Quality of life</strong></td>
<td></td>
</tr>
<tr>
<td>Functional Assessment of Cancer Therapy-Breast [37]</td>
<td>Assesses participants’ physical, social, family, emotional and functional well-being, and breast cancer–specific concerns</td>
</tr>
<tr>
<td><strong>SCT constructs</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Self-efficacy</strong></td>
<td></td>
</tr>
<tr>
<td>Exercise Self-Efficacy Scale [38]</td>
<td>Assesses beliefs in ability to be physically active over the next 12 weeks</td>
</tr>
<tr>
<td>Barriers Self-Efficacy Scale [38]</td>
<td>Assesses beliefs in ability to be physically active over the next 12 weeks despite common barriers</td>
</tr>
<tr>
<td><strong>Outcome expectations</strong></td>
<td></td>
</tr>
<tr>
<td>Multidimensional Outcome Expectations for Exercise Scale [39]</td>
<td>Assesses social, self-evaluative, and physical outcome expectations for physical activity</td>
</tr>
<tr>
<td><strong>Goal setting</strong></td>
<td></td>
</tr>
<tr>
<td>Exercise Goal-Setting Scale [40]</td>
<td>Assesses physical activity–related goal setting, self-monitoring, and problem solving</td>
</tr>
<tr>
<td><strong>Facilitators and barriers</strong></td>
<td></td>
</tr>
<tr>
<td>Social Support for Exercise Scale [41]</td>
<td>Measure physical activity support received from friends, family, and other survivors</td>
</tr>
<tr>
<td>Physical Activity Enjoyment Scale [42]</td>
<td>Measures enjoyment and satisfaction with current physical activity program</td>
</tr>
</tbody>
</table>

[b]SCT: social cognitive theory.

**Functional Performance**

Originally, the entire sample was used to complete functional performance measures in-person at each time point. However, as a result of the 2020 coronavirus global pandemic, these measures will only be completed on a subset (n=25) of the sample at both time points via a videoconference. Enrollment began in January 2020, and 25 women were enrolled in March 2020. These initial 25 participants completed in-person functional performance testing at baseline but were unable to complete 12-week functional performance testing because in-person research was suspended due to the coronavirus pandemic. Thus, no 12-week functional performance data were collected for the first 25 participants. However, functional performance data will be collected remotely from the next 25 participants to enroll.

Participants will complete the Short Physical Performance Battery (SPPB), a well-validated physical function performance measure [43,44]. The SPPB score is based on timed measures of gait speed, the ability to rise from a chair, and standing balance. Gait speed will be measured using the faster of 2 recorded times over a 4-m course. Chair stand time will be measured as the time needed to rise 5 times from a seated position in a chair, with arms folded across the chest. For the balance test, participants will be asked to maintain their feet side-by-side in semitandem and tandem positions for 10 seconds each. Each individual performance measure will be scored according to established cut-points [43] and aggregated for a total SPPB score. We will also use several items from the Senior Fitness Test [45], including the 8-foot up-and-go test, a test of physical agility and dynamic balance; the Arm Curl test, which assesses arm muscle strength endurance, specifically of the
biceps; and the 2-minute step test, an aerobic endurance test, which counts the number of full in-place steps completed in 2 minutes. Participants will also complete a 30-second one-leg stand test on each side (right and left), where they will be instructed to stand on 1 foot for up to 30 seconds and timed. Finally, participants will complete a 6-minute walk test [46] to assess their submaximal level of functional capacity.

Participants will be mailed a functional performance test tool kit that will include all of the supplies necessary to complete the functional performance tests. The study staff will administer functional performance testing via videoconference using a presentation with prerecorded videos detailing test set up and instructions. Each testing session will take approximately 60-90 minutes; 2-3 staff members certified in testing procedures will be present at the videoconference, varying depending on the testing time point. A staff member will administer the tests and be blinded to the condition; the other staff members will support the administrator and assist with any technology issues or in the event of an emergency. This staff member may or may not be blinded but will have no previous relationship (ie, not assigned to support coaching calls) with the participant. To ensure safety, participants will be asked to have another individual present at home during the test, if possible, and the study staff will be cardiopulmonary resuscitation or first-aid certified. In addition, the participant’s address and major cross streets will be confirmed and the nonemergency police number will be written in the participant file so that both study staff have access to this number in the event of an emergency during testing.

If the participant consents, the tests will be video recorded for quality control. Metrics on the feasibility of conducting these assessments via videoconference, including mailing issues and technology issues, will be recorded. Participants will be asked to complete a brief web-based survey to assess their experience with functional performance testing, including ease and satisfaction with using the videoconference software and functional performance test kit supplies, setting up testing, and following the written and videoconference instructions at each time point.

**Fidelity and Adherence**

Fidelity and adherence to each Fit2ThriveMB intervention and control group components are detailed in Textbox 1. All measures are objectively obtained unless otherwise noted. Coaching calls are recorded for all participants who consent to the audio recording.

**Textbox 1.** Measures of fidelity and adherence to intervention components.

<table>
<thead>
<tr>
<th>Fit2ThriveMB intervention</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Fitbit usage: average number of days of wearing the Fitbit (≥2000 steps/day) and proportion of days of wearing the Fitbit</td>
<td>• Cancer.Net usage: self-reported number of days using the app</td>
</tr>
<tr>
<td>• Fit2ThriveMB app: average number of days participants opened the app and average time spent using the app each day</td>
<td>• Support calls: total percentage of calls attended; average time per call; and fidelity to support the call script</td>
</tr>
<tr>
<td>• Daily goal setting: average number of days participants responded to the survey</td>
<td></td>
</tr>
<tr>
<td>• Fit2ThriveMB app social feed: total number of social feed posts; the average number of days of access to the social feed page</td>
<td></td>
</tr>
<tr>
<td>• Fit2ThriveMB app notifications: average number of messages read</td>
<td></td>
</tr>
<tr>
<td>• Fit2ThriveMB coaching calls: total percentage of coaching calls attended; average time per call; fidelity to coaching call script</td>
<td></td>
</tr>
</tbody>
</table>

**Additional Variables**

**Demographic and Disease Characteristics**

Participants will self-report on demographic characteristics and health status, including age, height, weight, race or ethnicity, marital status, education, income, and comorbid conditions. They will also self-report disease and treatment characteristics, including date of diagnosis, site of metastasis, hormone receptor status, and treatment received. The diagnosis and treatment data will be verified using medical records.

**SCT Constructs**

All SCT constructs will be assessed, including self-efficacy, goal setting, outcome expectations, and barriers and facilitators (Table 1).
Study Packet
All participants will be given a study packet following randomization. Individuals assigned to Fit2ThriveMB will be provided with a Fitbit, instructions on all intervention procedures, and instructions for setting up the Fitbit and downloading, setting up, and using the Fitbit and Fit2ThriveMB apps. Individuals assigned to the healthy lifestyle control groups will be given instructions on all intervention procedures and instructions for downloading, setting up, and using the Cancer.Net app.

Orientation
Participants will be scheduled for an in-person or videoconference orientation specific to the intervention condition they are assigned before starting the intervention. Study staff will reiterate the expectations for the condition to which they are assigned, answer any questions, and troubleshoot any technical issues they may be experiencing.

Intervention: Fit2ThriveMB Arm
The Fit2ThriveMB intervention arm consists of 3 intervention components described in detail below: the Fitbit, the Fit2ThriveMB app, and coaching calls. Participants will be encouraged to accumulate more steps throughout the day by (1) moving more and sitting less (ie, taking more steps and reducing the time they spend sedentary), (2) adding planned structured exercise sessions to their day, or (3) a combination of (1) and (2). They will also be provided with a more detailed exercise prescription if they decide to adopt approach (2) (Table 2). Participants will have access to the app for the duration of the 12-week study and the remaining duration of the funding period (up to approximately 12 months).

Table 2. Fit2ThriveMB exercise prescription.

<table>
<thead>
<tr>
<th>Week</th>
<th>Weekly exercise goal (min)</th>
<th>Number of sessions per week</th>
<th>Session duration (min)</th>
<th>Estimated steps per session</th>
<th>Rating of perceived exertion</th>
<th>Heart rate target (% of maximum heart rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>2-3</td>
<td>10-15</td>
<td>800-1200</td>
<td>1000-1500</td>
<td>9-12</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>2-3</td>
<td>15-20</td>
<td>1200-1600</td>
<td>1500-2000</td>
<td>9-12</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>3</td>
<td>20</td>
<td>1600</td>
<td>2000</td>
<td>9-12</td>
</tr>
<tr>
<td>4</td>
<td>75</td>
<td>3</td>
<td>25</td>
<td>2000</td>
<td>2500</td>
<td>9-12</td>
</tr>
<tr>
<td>5</td>
<td>90</td>
<td>3</td>
<td>30</td>
<td>2400</td>
<td>3000</td>
<td>9-12</td>
</tr>
<tr>
<td>6</td>
<td>105</td>
<td>3</td>
<td>35</td>
<td>2800</td>
<td>3500</td>
<td>9-12</td>
</tr>
<tr>
<td>7</td>
<td>120</td>
<td>3-4</td>
<td>30-40</td>
<td>2400-3200</td>
<td>3000-4000</td>
<td>9-12</td>
</tr>
<tr>
<td>8</td>
<td>135</td>
<td>3-4</td>
<td>30-45</td>
<td>2400-3600</td>
<td>3000-4500</td>
<td>9-15</td>
</tr>
<tr>
<td>9</td>
<td>150+</td>
<td>3-5</td>
<td>30-60</td>
<td>2400-4800</td>
<td>3000-6000</td>
<td>9-15</td>
</tr>
<tr>
<td>10</td>
<td>150+</td>
<td>3-5</td>
<td>30-60</td>
<td>2400-4800</td>
<td>3000-6000</td>
<td>9-15</td>
</tr>
<tr>
<td>11</td>
<td>150+</td>
<td>3-5</td>
<td>30-60</td>
<td>2400-4800</td>
<td>3000-6000</td>
<td>9-15</td>
</tr>
<tr>
<td>12</td>
<td>150+</td>
<td>3-5</td>
<td>30-60</td>
<td>2400-4800</td>
<td>3000-6000</td>
<td>9-15</td>
</tr>
</tbody>
</table>

aVigorous cells are empty in weeks 1-7 because vigorous activity is not recommended until week 8 to ensure proper progression.

Fitbit
Participants will be provided with a Fitbit InspireHR. They will be asked to download the Fitbit app and wear the Fitbit 24/7 throughout the 12-week study period. Fitbit measures PA intensity, steps, and heart rate and synchronizes directly with the Fitbit app, which automatically synchronizes with the Fit2ThriveMB app and provides Fitbit data to the study team in real time.

Fit2ThriveMB App
The custom-built Fit2ThriveMB app will encourage participants to increase their PA through SCT [22] principles, including goal setting, self-monitoring, personalized feedback on progress, app notifications, education on behavior change techniques, and social support. The Fit2ThriveMB app home screen and progress screen are shown in Figure 3. The app features are detailed below.
Daily Symptom Reporting and Goal Setting
Participants will be prompted each morning to report the intensity of their symptom burden on a scale from 1 to 5 by answering the following question: “To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?” [47]. On the basis of symptom rating, whether or not they met the previous day’s step goal and/or the number of steps achieved the previous day, participants will be provided with 3-4 goal options for that day (Figure 4). The goal algorithms are presented in Table 3. Standard goal logic options of increasing or decreasing steps by 10% or 20% or remaining constant will be applied to days where the steps of the previous days are between 3751 and 12,000. As the intervention’s goal is to increase steps, the algorithm never recommends less than 3000 steps, which is considered the floor value for absolutely defined moderate-intensity walking [48]. Special goal-setting logic is applied when the Fitbit is classified as not worn (ie, <2000 steps [49]) on the previous day, steps are low (ie, 2001-3750) on the previous day, and steps are high (>12,000) on the previous day.
If a participant fails to choose a goal, they are assigned the middle-level goal as their goal for the day. If a participant does not answer the symptom burden survey, their goal from the previous day is carried forward. The step goal-setting options for the first day of the intervention will be based on the average value of steps for all valid days of accelerometer data at baseline.
Table 3. Daily step goal logic.

<table>
<thead>
<tr>
<th>Symptom rating</th>
<th>Step goal option 1</th>
<th>Step goal option 2</th>
<th>Step goal option 3</th>
<th>Step goal option 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steps achieved: any value</td>
<td>No response</td>
<td>__^a</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Steps achieved: no data or &lt;2000 steps</td>
<td>1 or 2</td>
<td>Steps achieved in the last available day &gt;2000</td>
<td>Increase of 10%</td>
<td>Increase of 20%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Decrease of 10%</td>
<td>Steps achieved in the last available day &gt;2000</td>
<td>Increase of 10%</td>
</tr>
<tr>
<td></td>
<td>4 or 5</td>
<td>Decrease of 20%</td>
<td>Decrease of 10%</td>
<td>Steps achieved in the last available day &gt;2000</td>
</tr>
<tr>
<td>Steps achieved: goal met and 2000-3750 steps</td>
<td>1 or 2</td>
<td>3750</td>
<td>Increase of 10% (4125)</td>
<td>Increase of 20% (4500)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Decrease of 10% (3375)</td>
<td>3750</td>
<td>Increase of 10% (4125)</td>
</tr>
<tr>
<td></td>
<td>4 or 5</td>
<td>Decrease of 20% (3000)</td>
<td>Decrease of 10% (3375)</td>
<td>3750</td>
</tr>
<tr>
<td>Steps achieved: goal not met and 2000-3750 steps</td>
<td>1 or 2</td>
<td>Decrease of 10% (3375)</td>
<td>3750</td>
<td>Increase of 10% (4125)</td>
</tr>
<tr>
<td></td>
<td>3, 4, or 5</td>
<td>Decrease of 20% (3000)</td>
<td>Decrease of 10% (3375)</td>
<td>3750</td>
</tr>
<tr>
<td>Steps achieved: goal met and 3751-11,999 steps</td>
<td>1 or 2</td>
<td>Steps achieved in previous day</td>
<td>Increase of 10%</td>
<td>Increase of 20%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Decrease of 10%</td>
<td>Steps achieved in previous day</td>
<td>Increase of 10%</td>
</tr>
<tr>
<td></td>
<td>4 or 5</td>
<td>Decrease of 20%</td>
<td>Decrease of 10%</td>
<td>Steps achieved in previous day</td>
</tr>
<tr>
<td>Steps achieved: goal not met and 3751-11,999 steps</td>
<td>1 or 2</td>
<td>Decrease of 10%</td>
<td>Steps achieved in previous day</td>
<td>Increase of 10%</td>
</tr>
<tr>
<td></td>
<td>3, 4, or 5</td>
<td>Decrease of 20%</td>
<td>Decrease of 10%</td>
<td>Steps achieved in previous day</td>
</tr>
<tr>
<td>Steps achieved: goal met and ≥12,000 steps</td>
<td>1 or 2</td>
<td>Steps achieved in previous day</td>
<td>Increase of 10%</td>
<td>Increase of 20%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Decrease of 10%</td>
<td>Steps achieved in previous day</td>
<td>Increase of 10%</td>
</tr>
<tr>
<td></td>
<td>4 or 5</td>
<td>Decrease of 20%</td>
<td>Decrease of 10%</td>
<td>Steps achieved in previous day</td>
</tr>
<tr>
<td>Steps achieved: goal not met and 12,000+ steps</td>
<td>1 or 2</td>
<td>Decrease of 10%</td>
<td>Steps achieved in the previous day</td>
<td>Increase of 10%</td>
</tr>
<tr>
<td></td>
<td>3, 4, or 5</td>
<td>Decrease of 20%</td>
<td>Decrease of 10%</td>
<td>Steps achieved in the previous day</td>
</tr>
</tbody>
</table>

^aThe goal remains the same as the last option chosen or assigned.
^bN/A: not applicable.

**Self-monitoring**

Participants will receive feedback on their progress toward their daily step goal on the app home screen and weekly and monthly progress information in the fit progress section of the app. In addition, they can monitor progress toward their step goal on the Fitbit device screen.

**Fit Progress**

Patients will be provided with information on their steps, minutes of MVPA, and distance traveled in miles. They will be
able to view this information from the previous week, current week, and current month (ie, the previous 4 weeks) of the program.

1. Steps: a bar graph displays the steps achieved for the selected period. The number of steps reached on individual days of the week is listed in the graph if the previous week or current week’s time frame is selected. If the current month is selected, the average number of steps per week is displayed.

2. Minutes: a line graph displays minutes of MVPA achieved for the selected period. Minutes of MVPA achieved on individual days of the week will be listed under the graph if the previous week or current week’s time frame is selected. If the current month is selected, the total minutes of MVPA per week are displayed.

3. Distance: a line graph displays the distance traveled in miles for the selected period. The distance traveled on individual days of the week will be listed under the graph if the previous week or current week’s time frame is selected. If the current month is selected, the total miles per week are displayed.

**Fit Lessons**

Participants will be provided with educational information on PA and effective SCT behavior change strategies to incorporate more PA into their daily lives to increase their step count (eg, 10-minute lunchtime walk, parking further from entrances, and pacing while on the phone) and structured exercise via the app home screen.

**App Notifications**

The app will send participants text messages via app notifications, including tips to increase steps, a social digest, encouragement messages, real-time PA progress messages, and reminders to rate symptom burden, explore Fit Lessons, synchronize Fitbit, and engage in the social feed. A full list of message types, frequencies, and examples is provided in Table 4. Progress messages will be tailored to each patient’s fitness PA data. Participants will be instructed to enable notifications from the Fit2ThriveMB app during download. However, due to phone carrier companies’ restrictions, we cannot require that they enable this functionality. They can also change their app notification settings to disable these messages.
Participants will also be able to view their group members’ comment on other people’s posts, and share photos and website links. Participants will have the ability to write their own posts, include the following: hobbies, favorite physical activities, biography and optional picture. The social profile biographies participant will be given their own social profile, with a short randomized into groups using randomly permuted blocks. Each to provide encouragement and support. Participants will be communicate with 5 to 8 other participants in the intervention This will consist of a message board where participants can Social Feed

This will consist of a message board where participants can communicate with 5 to 8 other participants in the intervention to provide encouragement and support. Participants will be randomized into groups using randomly permuted blocks. Each participant will be given their own social profile, with a short biography and optional picture. The social profile biographies include the following: hobbies, favorite physical activities, motivation to become more active, goals for participating in the Fit2ThriveMB program, and favorite quote or mantra. Participants will have the ability to write their own posts, comment on other people’s posts, and share photos and website links. Participants will also be able to view their group members’ relative progress toward their daily step goals, as a circle around each patient’s profile picture will fill in as they accumulate steps each day (Figure 5). Participants will receive notifications when someone posts in the social feed, when they have been tagged in a post, and when someone in their group has met their activity goal for the day. A study team member will act as a moderator and closely monitor and facilitate social feed to enhance use and prevent misuse. The study staff will also use social feed as a way to provide important updates about the Fit2ThriveMB app and Fitbit, postdiscussion questions, and provide links to cancer survivorship and PA educational resources (Textbox 2 for examples). During weeks 1-7 of the intervention, there will be 2 preplanned posts in the social feed. During weeks 8-12, the study staff will post 1 preplanned post a week in the social feed. Throughout the intervention, study staff will also post important updates regarding study information and logistics as needed, which may increase the total number of posts per week.

Table 4. Fit2ThriveMB app notification types and examples.

<table>
<thead>
<tr>
<th>Message type</th>
<th>Description</th>
<th>Frequency</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptom message</td>
<td>Reminder to complete symptom burden questionnaire</td>
<td>Daily within 1 hour of wake time</td>
<td>• “Time to answer your daily survey and pick today’s step goal!”</td>
</tr>
<tr>
<td>Symptom reminder</td>
<td>3 reminders every 15 minutes to complete a survey</td>
<td>Every morning if the participant does not complete after the first message</td>
<td>• “Reminder: complete your symptom burden survey and choose today’s step goal!”</td>
</tr>
<tr>
<td>Social cognitive theory behavior change</td>
<td>Message to remind participants to look at the educational materials to learn more about physical activity and healthy habits</td>
<td>1 time per week</td>
<td>• “Did you know keeping a record of your activity and reviewing it can make you much more likely to reach your goals? Read more about daily tracking on the F2TMB Fit Lessons!”</td>
</tr>
<tr>
<td>Increasing steps</td>
<td>Educational message with ideas on how to increase daily step count by incorporating increasing steps into the daily routine</td>
<td>2 messages per week</td>
<td>• “Take the stairs instead of the elevator today. Small changes like this can count toward your daily step goal!”</td>
</tr>
<tr>
<td>Rescue</td>
<td>Provides motivation to get in some extra steps before the day ends and reach the goal. Messages are stratified based on if they have ≥2500 steps</td>
<td>Daily if participant has not achieved step goal by 4 PM</td>
<td>• “Almost there, you have [X] steps to go until you reach today’s goal. You can do it!”</td>
</tr>
<tr>
<td>Sync reminder</td>
<td>Reminder to sync Fitbit or add any activity for the day</td>
<td>Daily if the participant does not have data 1 hour before bedtime</td>
<td>• “Remember to sync your Fitbit data and enter any other activities you did today!”</td>
</tr>
<tr>
<td>Social goal</td>
<td>Any participant in the group reached their step goal for that day</td>
<td>Up to 8 times a day (depends on groups size) if someone in the groups meets their goal</td>
<td>• “[Z] reached 100% of her step goal today! Head to your newfeed to congratulate her!”</td>
</tr>
<tr>
<td>Social Digest</td>
<td>Message consisting of a number of new posts, comments to own posts, or tags from the social digest</td>
<td>Every 30 minutes, if something new is posted</td>
<td>• “There are [X] new posts and you’ve been mentioned [Y] times in your Fit2ThriveMB group!”</td>
</tr>
<tr>
<td>Reengagement</td>
<td>When a participant has not posted or commented in the social feed</td>
<td>Every 4 days, if the participant has not posted</td>
<td>• “Your Fit2ThriveMB group members miss you! Let them know how you are doing.”</td>
</tr>
</tbody>
</table>
| Goal progress                 | Provides encouraging message once 50% of the daily goal is met               | Encouraging messages when transmitted Fitbit data indicate they reached 50%, 75%, and 100% of the daily step goal | • “You’ve hit 50% of your step goal for the day, keep up the good work!”
|                               |                                                                             |                            | • “Way to hit 75% of your daily goal! Only [X] steps left, let’s do this.”
|                               |                                                                             |                            | • “100% of today’s step goal accomplished, well done!” |

**Social Feed**

This will consist of a message board where participants can communicate with 5 to 8 other participants in the intervention to provide encouragement and support. Participants will be randomized into groups using randomly permuted blocks. Each participant will be given their own social profile, with a short biography and optional picture. The social profile biographies include the following: hobbies, favorite physical activities, motivation to become more active, goals for participating in the Fit2ThriveMB program, and favorite quote or mantra. Participants will have the ability to write their own posts, comment on other people’s posts, and share photos and website links. Participants will also be able to view their group members’ relative progress toward their daily step goals, as a circle around each patient’s profile picture will fill in as they accumulate steps each day (Figure 5). Participants will receive notifications when someone posts in the social feed, when they have been tagged in a post, and when someone in their group has met their activity goal for the day. A study team member will act as a moderator and closely monitor and facilitate social feed to enhance use and prevent misuse. The study staff will also use social feed as a way to provide important updates about the Fit2ThriveMB app and Fitbit, postdiscussion questions, and provide links to cancer survivorship and PA educational resources (Textbox 2 for examples). During weeks 1-7 of the intervention, there will be 2 preplanned posts in the social feed. During weeks 8-12, the study staff will post 1 preplanned post a week in the social feed. Throughout the intervention, study staff will also post important updates regarding study information and logistics as needed, which may increase the total number of posts per week.
The total time spent moderating and posting by the study staff will be approximately 30-40 minutes per week.

**Figure 5.** Fit2ThriveMB app social feed features.

**Textbox 2.** Example Fit2ThriveMB social feed moderator posts.

**Physical activity discussion**

- “What is your biggest barrier to becoming physically active? What are you trying to do to overcome this barrier? Reply to this post, and if you like another group member’s response, "favorite" it by clicking on the star below!”

- “What are some of your favorite activities to get your daily steps in? Reply to this post and share with the group so everyone has some new ideas.”

**Physical activity information and survivorship news and resources**

- “If you’re trying to find new, different ways to get your steps in, click on the link below for some ideas as a jumping off point. Click on the "star" below to favorite this post if you are looking for new ideas or reply with some others that work for you!”

- “If you are looking for events in the Chicago area, Susan G Komen has the following list of events for the 2020 year so far. "Favorite" this post if you are interested in attending one of these events or reply with other event ideas!”

**Study information and logistics**

- “Your week 3 assessment questionnaire was sent out today! Please complete the questionnaire as soon as possible.”

- “Activity monitors for your week 12 assessments will be mailed tomorrow. Keep an eye on your mail for the assessment packet. Keep up the great work in the homestretch!”

**Coaching Calls**

Participants will receive weekly coaching calls that will provide feedback on the previous week’s symptom burden and progress on PA goals, review personalized goals and strategies for increasing PA for the next week, inquire about any issues or injuries, and cover at least 1 SCT [22] behavior change strategy (Table 5 for topics covered). Participants will be instructed to call or email their coaches throughout the week if they encounter...
any issues. The coaching calls will be semistructured and recorded to ensure fidelity.

Table 5. Weekly coaching call topics for Fit2ThriveMB intervention group.

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intro to coaching calls, getting started, and effective goal setting</td>
</tr>
<tr>
<td>2</td>
<td>Benefits of increasing physical activity and increasing steps</td>
</tr>
<tr>
<td>3</td>
<td>Benefits of reducing sitting or overcoming barriers to reducing sedentary behavior</td>
</tr>
<tr>
<td>4</td>
<td>Finding activities you enjoy</td>
</tr>
<tr>
<td>5</td>
<td>Self-monitoring</td>
</tr>
<tr>
<td>6</td>
<td>Managing symptoms</td>
</tr>
<tr>
<td>7</td>
<td>Overcoming barriers</td>
</tr>
<tr>
<td>8</td>
<td>Increasing self-efficacy</td>
</tr>
<tr>
<td>9</td>
<td>Realistic outcome expectations</td>
</tr>
<tr>
<td>10</td>
<td>Social support</td>
</tr>
<tr>
<td>11</td>
<td>Healthy rewards</td>
</tr>
<tr>
<td>12</td>
<td>Relapse prevention</td>
</tr>
</tbody>
</table>

**Technical Support**

Participants will be provided with detailed instructions for setting up their Fitbit and Fit2ThriveMB app. Participants will be encouraged to contact study staff via phone, email, or the Fit2ThriveMB app social feed as soon as possible if any technical issues arise. All participants will be instructed to inform the study team immediately if they lose their Fitbit or if it malfunctions, and they will be sent a new device. They will also be asked about technical issues during their weekly coaching calls, which the study team will take note of, discuss with the technical support team, and follow up with the participant if needed.

**Privacy and Confidentiality**

Participants will use a Quick Response code specific to them to log into the Fit2ThriveMB app. To protect their privacy, no personal information is stored or associated with their app account. All Fitbit and app data will be collected using the University server cluster, which has limited physical access, is firewalled, and regularly monitored for security issues. All phone encrypted data transmissions use a secure sockets layer protocol with a unique token for each participant. All data are backed up regularly, and any personal health information and Health Insurance Portability and Accountability Act data are stored on separate data clusters with unique keys and limited firewalled access. Participants will be informed of all potential privacy and confidentiality risks and practices in place to ensure protection in the informed consent form.

**Waitlist Attention Control: Healthy Lifestyle Arm**

Participants assigned to this condition will be instructed to go about their usual activities. They will be provided with instructions for downloading and using the Cancer.Net app and encouraged to regularly use the Cancer.Net app features, including information about health and well-being, treatment guidelines specific to their cancer type, symptom tracking, and appointment tracking. Although the Cancer.Net app does provide content related to PA, this information is not prominent, and no information related to PA will be covered in this condition. To match Fit2ThriveMB contacts, participants will also receive weekly calls that will last approximately 10-15 minutes and cover health and well-being topics (Table 6) and be directed to content within the Cancer.Net app related to each topic. They will be provided with the Fitbit and Fit2ThriveMB app and all of its features, including the social feed, after completion of week 12 assessments and provided with a start date to begin receipt of the 12-week program. They will have the option to attend a brief program orientation with study staff and access to the Fit2ThriveMB app for the duration of the funding period (up to approximately 12 months).
Table 6. Weekly healthy lifestyle control group call topics.

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mindfulness</td>
</tr>
<tr>
<td>2</td>
<td>Social support</td>
</tr>
<tr>
<td>3</td>
<td>Symptom tracking</td>
</tr>
<tr>
<td>4</td>
<td>Art therapy and doing and finding activities you enjoy</td>
</tr>
<tr>
<td>5</td>
<td>Sleep</td>
</tr>
<tr>
<td>6</td>
<td>Fatigue</td>
</tr>
<tr>
<td>7</td>
<td>Cognition</td>
</tr>
<tr>
<td>8</td>
<td>Hydration</td>
</tr>
<tr>
<td>9</td>
<td>Nutrition</td>
</tr>
<tr>
<td>10</td>
<td>Taste changes</td>
</tr>
<tr>
<td>11</td>
<td>Healthy grocery shopping</td>
</tr>
<tr>
<td>12</td>
<td>Stress management</td>
</tr>
</tbody>
</table>

Adverse Event and Safety Monitoring

Safety will be measured based on the number and severity of adverse events reported. All participants will be instructed to report any adverse events to the study staff within 24 hours of the event. Participants will answer nonspecific adverse event questions at weeks 4, 8, and 12 to collect any information on any adverse events that may have occurred but were not reported to the study team.

Data Analytic Plan

Power

Sample size calculations (n=50) were based on the successful completion rate defined as an attrition rate of ≤20% (ie, complete 12-week assessment) and adherence to daily goals and wearing Fitbit ≥70% (59/84) of study days. With half of the participants (n=25) assigned to each condition, there is >85% power to differentiate between a 60% (control) and 80% (Fit2ThriveMB) successful completion rate with a one-tailed, exact test for binomial proportion at a .05 significance level (SAS version 9.4).

Aim 1: Feasibility and Acceptability

Feasibility and acceptability will be analyzed using descriptive statistics (frequencies, means, and SDs). Feasibility will primarily be evaluated using the successful completion rate defined as ≥80% (40/50) of participants who enroll in the intervention will remain in the intervention and participants will meet their daily goal and wear Fitbit ≥70% (59/84) of study days. If a successful completion rate is achieved, Fit2ThriveMB will be deemed feasible. Additional measures of feasibility, fidelity, acceptability, and safety will also be summarized and considered with the successful completion rate before proceeding with a more definitive trial.

Aims 2 and 3: Examine Outcome Patterns for PA, Symptoms, Functional Performance, and QoL

We will employ an intent-to-treat approach by including all participants recruited in the study, regardless of compliance. Every effort will be made to collect all outcome measures, even if a participant does not engage in assigned treatments. Analysis of covariance will be used to compare study groups concerning mean changes in each outcome from baseline to 12 weeks with baseline adjustment. Data from these analyses will provide estimated means, SEs, and preliminary effect sizes for the Fit2ThriveMB intervention. These data will be used to identify a primary end point and further refine the Fit2ThriveMB intervention to be tested in a more definitive trial. We will also use regression models to explore the relationship between changes in each outcome and intervention adherence. This study is not designed to draw conclusions about Fit2ThriveMB efficacy without further study.

Results

This study was funded in March 2019. Data collection began in February 2020 and ended in February 2021. A total of 49 women were randomized, and the results are expected to be published in the summer of 2021. Preliminary data indicate that the sample is 86% (n=42) White, 6% (n=3) African American, 6% (n=2) Asian, 4% (n=2) unknown (participants did not want to disclose their race), and 2% (n=1) other; 10% (n=5) of the sample identified as Hispanic or Latina.

Discussion

Principal Findings

Advances in the treatment of metastatic breast cancer have resulted in a growing population of women living with metastatic breast cancer [1]. However, the addition of years to life does not necessarily translate to quality years. Increasing PA is a potentially modifiable behavior that could attenuate the adverse physical and psychological health effects of living with metastatic breast cancer. Very few PA promotion interventions have been designed and tested in patients with metastatic breast cancer. In fact, most trials specifically exclude metastatic patients. Furthermore, many of the published trials on PA for patients with metastatic breast cancer require at least some
on-site exercise sessions and have required significant adjustments to the planned exercise dose (ie, frequency and intensity) to accommodate the needs of metastatic patients. Thus, the feasibility and generalizability of these programs are limited.

**Limitations and Strengths**

The Fit2ThriveMB study has several limitations. Our sample may not be representative of all patients with metastatic breast cancer because of the recruitment strategies employed and the requirements to use technology. Participants in this trial are likely to be more technologically savvy, motivated to change their PA behavior, and potentially healthier than the general patient population with metastatic breast cancer. Recruitment from a large academic medical center may also bias the sample in favor of White, more affluent, and highly educated participants. We actively recruited and enrolled diverse samples. However, diversity should be enhanced in the future, with larger studies by recruiting more broadly in community cancer centers, community organizations, and support groups. Another concern may be the potential for low intervention adherence. To prevent adherence problems, the significance of the project and participation expectations will be clearly stated during recruitment and orientation, which will be incentivized for the completion of assessments. Participants will be called weekly, reminders will be sent for all study-related milestones, and participants in the intervention group will be contacted if they do not wear the Fitbit for ≥7 days. In addition, we anticipate that the highly tailored nature of the intervention will significantly reduce attrition and increase adherence. We do not limit participation in this study based on the treatment status or characteristics that could impact activity levels. We anticipate that these variables will not differ between groups due to randomization, but we will conduct analyses to confirm this. If not, we will explore the potential impact of these factors on trial outcomes. Finally, we did not include a true no-contact control condition. Accumulated evidence indicates that remaining sedentary is unlikely to have any positive effect on our outcomes of interest in low-active patients with cancer. Our design is such that we will be able to compare the intervention condition with a control group that receives similar attention and an app to use but no PA promotion resources, which will allow us to delineate the effects of the intervention itself from the attention received from the study staff.

The Fit2ThriveMB trial was designed to determine the feasibility and acceptability of a 12-week technology-based PA promotion intervention in patients with metastatic breast cancer. We will also explore outcome patterns suggesting the efficacy of Fit2ThriveMB on PA, symptom burden, QoL, and functional performance. Despite the limitations noted above, the Fit2ThriveMB intervention has many potential advantages. First, this research will provide a better understanding of how to effectively promote PA in patients with metastatic breast cancer, an understudied population. Second, this is the first study to focus on increasing the cumulative daily PA of any intensity in patients with metastatic breast cancer. Previous studies have focused on high doses of more intense MVPA, which is not feasible for many patients with metastatic breast cancer. Light-intensity activity has been associated with health benefits, including the prevention of functional decline [10] and mortality [11] and improved QoL [11,12] and body composition [13]. This approach may have substantial health benefits for patients with metastatic breast cancer, facilitate gradual and safe PA adoption by enhancing mastery experiences, and be a more feasible behavioral target than high volumes of more strenuous activity [50,51]. Third, this is the first study to test a theory-based, technology-based PA promotion intervention in patients with metastatic breast cancer using the technology they already own. This approach allows for remote monitoring of patients and prevents many of the barriers associated with on-site programs (eg, travel, time constraints, or difficulties parking). Finally, this is the first study to adopt a highly tailored and individualized approach that takes into account the PA level and symptom burden in any patient population with cancer. This will likely increase adherence and uptake of the intervention by accounting for each survivor's capabilities.

**Conclusions**

On the basis of the results of this study, we will determine the feasibility and acceptability of the Fit2ThriveMB intervention and examine outcome patterns that suggest efficacy. If the results indicate that Fit2ThriveMB is feasible and likely efficacious, the intervention will be refined to enhance usability and reflect changes in technology. We will then conduct a large-scale, fully powered trial of the entire Fit2ThriveMB intervention package on changes or maintenance of outcomes examined in this study and additional outcomes (eg, biomarker end points and fitness) in additional populations with advanced cancer. Ultimately, the proposed study will provide key evidence to support the feasibility, acceptability, and health benefits of increasing PA in patients with metastatic breast cancer using a scalable intervention strategy that could be easily integrated into care to improve health and disease outcomes.

**Acknowledgments**

This work was supported by the National Institutes of Health R21CA239130 and K07CA196840 awarded to SP. LA is supported by T32CA193193, and WW is supported by R37CA225877.

**Conflicts of Interest**

None declared.

**References**

https://www.researchprotocols.org/2021/4/e24254

JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e24254 | p.132

(page number not for citation purposes)


**Abbreviations**

**mHealth**: mobile health  
**MVPA**: moderate-to-vigorous physical activity  
**PA**: physical activity  
**QoL**: quality of life  
**RCT**: randomized controlled trial  
**SCT**: social cognitive theory  
**SPPB**: Short Physical Performance Battery

©Siobhan Phillips, Payton Solk, Whitney Welch, Lisa Auster-Gussman, Marilyn Lu, Erin Cullather, Emily Torre, Madelyn Whitaker, Emily Izenman, Jennifer La, Jungwha Lee, Bonnie Spring, William Gradishar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
A Live Video Program to Prevent Chronic Pain and Disability in At-Risk Adults With Acute Orthopedic Injuries (Toolkit for Optimal Recovery): Protocol for a Multisite Feasibility Study

Ana-Maria Vranceanu1,2, PhD; Jafar Bakhshaie1,2, MD, PhD; Mira Reichman1, BA; James Doorley1,2, MA; A Rani Elwy1,4, PhD; Cale Jacobs5, PhD; Neal Chen6, MD; John Esposito6, MD; David Laverty7, MD; Paul E Matuszewski5, MD; Amirreza Fatehi7, MD; Lucy C Bowers7, BSc; Mitchel Harris6, MD; David Ring7, MD

1Integrated Brain Health Clinical and Research Program, Massachusetts General Hospital, Boston, MA, United States
2Harvard Medical School, Boston, MA, United States
3Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, Providence, RI, United States
4Center for Healthcare Organization and Implementation Research, VA Bedford Healthcare System, Bedford, MA, United States
5Department of Orthopaedic Surgery & Sports Medicine, College of Medicine, University of Kentucky, Lexington, KY, United States
6Department of Orthopaedic Surgery, Massachusetts General Hospital, Boston, MA, United States
7Department of Surgery and Perioperative Care, Dell Medical School, The University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Ana-Maria Vranceanu, PhD
Integrated Brain Health Clinical and Research Program
Massachusetts General Hospital
1 Bowdoin Square, Suite 100
Boston, MA, 02114
United States
Phone: 1 6177244977
Email: avranceanu@mgh.harvard.edu

Abstract

Background: Despite the pivotal role of psychosocial factors in pain and disability after orthopedic injury, there are no evidence-based preventive interventions targeting psychosocial factors in patients with acute orthopedic injuries. We developed the first mind-body intervention focused on optimizing recovery and improving pain and disability in patients with acute orthopedic injuries who exhibit high levels of catastrophic thinking about pain and/or pain anxiety (Toolkit for Optimal Recovery [TOR] after orthopedic injury). In a pilot single-site randomized controlled trial (RCT), the TOR met a priori set benchmarks for feasibility, acceptability, and satisfaction. The next step in developing TOR is to conduct a multisite feasibility RCT to set the stage for a scientifically rigorous hybrid efficacy-effectiveness trial.

Objective: The objective of this study is to conduct a rigorous multisite feasibility RCT of TOR to determine whether the intervention and study methodology meet a priori set benchmarks necessary for the successful implementation of a future multisite hybrid efficacy-effectiveness trial. In this paper, we describe the study design, manualized treatments, and specific strategies used to conduct this multisite feasibility RCT investigation.

Methods: This study will be conducted at 3 geographically diverse level 1 trauma centers, anonymized as sites A, B, and C. We will conduct a multisite feasibility RCT of TOR versus the minimally enhanced usual care (MEUC) control (60 patients per site; 30 per arm) targeting a priori set feasibility benchmarks. Adult patients with acute orthopedic injuries who endorse high pain catastrophizing or pain anxiety will be recruited approximately 1-2 months after injury or surgery (baseline). Participants randomized to the TOR will receive a 4-session mind-body treatment delivered via a secure live video by trained clinical psychologists. Participants randomized to the MEUC will receive an educational booklet. Primary outcomes include feasibility of recruitment, appropriateness, feasibility of data collection, acceptability of TOR (adherence to sessions), and treatment satisfaction across all sites. We will also collect data on secondary implementation outcomes, as well as pain severity, physical and emotional function, coping skills, and adverse events. Outcomes will be assessed at baseline, posttreatment, and at the 3-month follow-up.

Results: Enrollment for the RCT is estimated to begin in June 2021. The target date of completion of the feasibility RCT is April 2024. The institutional review board approval has been obtained (January 2020).
Conclusions: This investigation examines the multisite feasibility of TOR administered via live videoconferencing in adult patients with acute orthopedic injuries. If feasible, the next step is a multisite, hybrid efficacy-effectiveness trial of TOR versus MEUC. Preventive psychosocial interventions can provide a new way to improve patient and provider satisfaction and decrease suffering and health care costs among patients with orthopedic injuries who are at risk for chronic pain and disability.

International Registered Report Identifier (IRRID): PRR1-10.2196/28155

KEYWORDS
orthopedic; musculoskeletal; prevention; chronic pain; disability; intervention; video; telehealth; mobile phone

Introduction

Background

Acute orthopedic injuries, such as fractures and dislocations, are highly prevalent and costly [1-3]. Regardless of objective healing (eg, via radiography or magnetic resonance imaging), up to 50% of patients with acute orthopedic injuries will eventually develop chronic pain and disability requiring additional medical treatments, further burdening the health care system [1]. During the acute phase of orthopedic injury, psychosocial factors such as pain anxiety and catastrophizing (ie, intensely negative and inaccurate thoughts regarding pain and activity) are among the strongest predictors of pain and disability [4-8], regardless of pain severity [9], pain location [7,10], or injury type [6,11,12]. To prevent chronic pain and disability among patients with orthopedic injuries, early detection and targeted intervention focused on these psychosocial factors is crucial [13,14].

Currently, there are no evidence-based preventive interventions targeting psychological factors in patients with acute orthopedic injuries [15]. Indeed, the model of care for acute orthopedic injuries is rigidly confined to the biomedical paradigm, relying almost exclusively on surgical intervention and pain medication. Although recent attention to psychosocial factors in orthopedic populations has increased surgeons’ awareness of the clinical relevance of psychological factors [16], surgeons still report reluctance and discomfort when discussing these concerns with patients or making referrals to psychological service providers [16-18]. Typically, referrals for psychological services are made when orthopedic patients have already developed chronic pain and disability, missing the critical period for preventive intervention [17,18]. Other barriers to the provision of psychosocial care for patients with orthopedic injuries include stigma (eg, related to diagnostic labels and mental health treatment), transportation and time constraints, and lack of psychological care providers with training relevant to the orthopedic population [5,19].

Mind-body interventions tailored to address the psychosocial needs of patients with acute orthopedic injuries offer a novel approach for improving pain and disability [20-25]. Mind-body interventions have become increasingly popular in recent years among medical patients, including those with orthopedic conditions, and represent a powerful avenue for engaging patients in psychosocial care despite the stigma associated with mental health services [20-25]. In line with this trend, our group developed Toolkit for Optimal Recovery (TOR) after orthopedic injury, the first mind-body program focused on optimizing recovery and improving pain and disability-related outcomes in patients with acute orthopedic injuries who are at risk for chronic pain and disability as a function of psychological factors (ie, exhibit high levels of catastrophic thinking about pain and/or pain anxiety) [26]. Rooted in the theoretical framework of the fear avoidance model [27], TOR targets catastrophic thinking about pain (eg, misconceptions about pain and activity, hopelessness, helplessness, magnification of pain) and pain anxiety (eg, negative pain-related thoughts, fear of pain, hypervigilance, heightened physiological reactivity to pain sensations, and pain-related escape and avoidance).

TOR helps patients confront rather than avoid their pain-related experiences through teaching relaxation response skills (ie, exercises designed to decrease heart and breath rate) [28] and mindfulness skills (ie, intentional self-regulation of attention from moment to moment) [29], correction of misconceptions about pain and physical activity, acceptance (eg, of pain), values-based activity engagement, and activity pacing. Conceptually, participation in the TOR will be associated with decreased catastrophic thinking about pain and pain anxiety. In turn, these factors will lead to improved pain and disability-related outcomes and reduce the risk of chronic pain and disability (see Figure 1 for an illustration of this conceptual model). TOR was developed iteratively with feedback from patients, surgeons, and other orthopedic care providers and uses a telehealth delivery method with the goal of increasing access, cost-effectiveness, and service satisfaction [26].
We previously found evidence for feasibility, acceptability, and satisfaction for TOR in a pilot feasibility randomized controlled trial (RCT) conducted at a single orthopedic trauma center [26]. Despite these positive outcomes, our pilot RCT revealed several methodological weaknesses that needed to be addressed before a hybrid efficacy-effectiveness trial. In addition to generalizability concerns due to single-site sampling, half of the surgeons within the recruiting orthopedic department (2 out of the 4) did not cooperate in making referrals. Specifically, surgeons reported concerns about disrupting clinical flow within the orthopedic department and expressed skepticism regarding the clinical relevance of psychological interventions for patients with acute orthopedic injuries. These challenges were consistent with previous work demonstrating surgeons’ discomfort referring patients to psychological interventions because of factors such as low mental health literacy, lack of training, lack of confidence in the system, and fear of upsetting patients [16].

Despite surgeons’ reluctance to refer patients to our pilot RCT, we did observe that a warm handoff (where surgeons personally introduced research assistants [RAs] to patients) was the most effective manner to recruit and enroll participants. We now understand that buy-in from all surgeons and orthopedic staff is pivotal for successfully implementing our study and interventions. To increase buy-in, we need to educate orthopedic care teams on the relevance of psychosocial care. We also need to listen to their feedback on how to seamlessly integrate our study into clinical flow to ensure that patient care is optimized and not disrupted. The study also highlighted the importance of recruiting a racial, ethnic, and economically diverse sample to increase the generalizability of results.

To address these limitations and to maximize the success of our planned multisite feasibility RCT, we conducted focus groups with orthopedic surgeons and staff at 3 level 1 trauma centers (86 participants recruited from the 3 sites). We used elements of Proctor’s Implementation Outcomes Framework [30] and the Consolidated Framework for Implementation Research [31] to understand the (1) barriers and facilitators to integrate psychosocial care into orthopedic care broadly, (2) barriers and facilitators to implement our study interventions and procedures, and (3) feedback for the development of training materials to educate orthopedic providers regarding the importance of psychosocial factors and facilitate provider buy-in within each site. Using this method, we gathered information necessary to optimize the implementation of TOR and related study procedures using the implementation strategies of the Expert Recommendations for Implementing Change group [32], which helped refine our protocol.

**Objectives**

Here, we outline the protocol for our multisite feasibility RCT (informed and refined by the findings of the focus groups), which is a necessary step in preparing for a future hybrid efficacy-effectiveness trial. The goal of this trial is to evaluate the feasibility of our study procedures and interventions (TOR and minimally enhanced usual care [MEUC]) at 3 sites based on a priori set benchmarks of acceptability, feasibility, appropriateness, and fidelity (at the provider and patient levels). Our secondary goal is to maximize the recruitment and retention of racial and ethnic minorities to achieve generalizable results. The results will inform a fully powered, multisite, hybrid efficacy-effectiveness study of TOR versus MEUC.

**Methods**

**Study Design**

This study will use a multisite RCT design to evaluate the feasibility of TOR delivered via secure live video, in line with the guidelines for iterative optimization of study design and methodology before conducting a fully powered, hybrid efficacy-effectiveness trial. The institutional review board (IRB) at site A approved all the study procedures. With a cross-site reliance agreement, the site A IRB and an external data safety and monitoring board will oversee the implementation of all study-related procedures at all sites.

The study protocol was informed by and refined based on the qualitative data collected during the focus groups and exit interviews. Specifically, we revised the initially proposed study protocol for the multisite RCT, including milestones, recruitment and retention procedures, and inclusion and exclusion criteria and added strategies to maximize success based on anticipated potential challenges at each site (see Multimedia Appendix 1 for peer review of this research proposal by the National Institutes of Health).
Setting

This multisite study will be conducted at 3 geographically diverse level 1 trauma centers, anonymized as sites A, B, and C.

Inclusion and Exclusion Criteria

Participants will be patients from sites A, B, and C, with acute orthopedic musculoskeletal injuries who are at risk for chronic pain and disability based on high pain catastrophizing or anxiety scores (see the Assessments section) and who meet other inclusion and exclusion criteria (Textbox 1). At each of the 3 sites, we will recruit and enroll approximately 60 patients with acute musculoskeletal injuries (1-2 months after injury or surgery; 180 participants in total), representing a target sample size in accordance with the guidelines for feasibility trials.

Textbox 1. Inclusion and exclusion criteria for participants in the study.

**Inclusion criteria:**
- Male and female outpatients in a level 1 trauma center of one of the 3 sites, aged 18 years or above
- Sustained a single acute orthopedic injury (eg, fracture, dislocation, or rupture) approximately 1 to 2 months earlier (acute phase)
- Scored ≥20 on the Pain Catastrophizing Scale [31] or ≥40 on the Pain Anxiety Symptoms Scale–Short Form [32]
- Willing to participate and comply with the requirements of the study protocol, including randomization, questionnaire completion, and potential home practice and weekly sessions
- Free of concurrent psychotropic medication for at least 2 weeks before initiation of treatment or stable on current psychotropic medication for a minimum of 6 weeks and willing to maintain a stable dose (ie, no psychotropics or stable for >6 weeks)
- Cleared by the orthopedic surgeon for study participation
- Able to meaningfully participate (eg, speak English and have a stable living situation as determined by the medical staff at each site)

**Exclusion criteria:**
- Diagnosed medical illness that is expected to worsen in the next 3 months (eg, malignancy)
- Serious mental illness or instability for which hospitalization may be likely in the next 3 months
- Current suicidal ideation
- Other serious injuries that occurred alongside the orthopedic injury
- Lifetime history of schizophrenia, bipolar disorder, or other psychotic disorder
- Current substance use disorder
- Practice of yoga, meditation, or other mind-body techniques once per week for 45 minutes or more within the past 3 months
- Currently in litigation or under workman’s comp
- Surgery complications (eg, infection or need for repeat surgery)
- Self-reported pregnancy

Identifying, Recruiting, Consent, and Randomization Procedures

The RA at each site will review the electronic medical records of the outpatient orthopedic clinics on a daily basis, or when instructed by the medical or surgical staff, to identify participants coming into the clinic who have had an injury or a surgery approximately 1 to 2 months earlier. All sites have routine follow-up appointments scheduled during this timeframe.

The RA will provide a list of potentially eligible patients to the orthopedic medical staff, and the medical assistant in charge of checking patients into rooms will inform the patients that they are potential study participants, obtain verbal consent from them to participate in the screening procedures, and ask the patients to complete the self-report measure of pain catastrophizing and pain anxiety [33,34]. If there are too many eligible patients in one day, we will prioritize racial and ethnic minorities. For patients who screen in on either of these 2 measures, the medical assistant or RA will notify the orthopedic surgeon that the patient is a potential study participant (verbally or via a note with the study logo attached to the door). Surgeons will also receive information regarding the race and ethnicity of each participant collected by the RA from the electronic medical records in order to prioritize study referrals. The orthopedic surgeon will perform the medical visit and subsequently introduce the study to the potential participant using a predetermined script and materials and procedures developed with information from the focus groups. For patients who express interest, the surgeon will conduct a warm handoff referral to the RA, who will finish the screening process. If patients across sites are consistently unable to stay past their scheduled appointment window to finish all baseline procedures, we will flexibly utilize an alternate recruitment procedure in which patients are screened by phone before coming in for their 1- to 2-month follow-up visit so that they can be notified in advance if they are eligible and budget time for the research visit.
If a participant meets the eligibility criteria, the RA will meet with the participant in a private location following their clinic visit to describe the study in detail, including the consent form. All patients will have the opportunity to ask questions and will be given time to consider whether to participate. Participants who choose to participate will be asked to sign the consent form. Participants who choose to participate will be asked to sign the consent form. Participants will receive a copy of the consent form. Next, participants will complete the baseline assessments in the clinic and will be randomized. A trained RA will assist with performance-based measures of physical function (the grip test will be used for participants with upper extremity injuries or walk tests for participants with lower extremity injuries). Self-report measures will be completed using an iPad (Apple Inc). The RA will be available to ensure that all questions are answered and assist with completion, as needed. Participants will be compensated after baseline assessments.

After baseline assessments, participants will be randomized to TOR or MEUC in a 1:1 ratio according to a computer-generated randomization schedule. Participants randomized to the TOR will receive instructions for live video intervention before leaving the clinic. This includes (1) installing Zoom (Zoom Video Communications) and teaching participants how to use it, (2) scheduling their 4 weekly intervention sessions with the clinician, (3) setting up EZ Texting to receive reminders for sessions and home practice, and (4) installing the TOR web platform (with session content and guided exercises) as a smartphone app. Participants who do not own a smartphone will be given one, and the RA will set up their phone at the first visit. We will also provide study participants with a data plan for the duration of the study or brainstorm ways to access free internet (eg, library or friends’ houses), if needed.

The RA will provide a printed educational booklet along with access to a web-based version of the booklet content to participants randomized to the MEUC. Participants will be assisted in accessing this website from their preferred device and bookmarking the website, as desired. Participants in MEUC who do not own a smartphone will not be provided with one, as they can also access the educational information in printed form.

**TOR Program Structure**

TOR is a 4-session live video–based, individual mind-body program that aims to optimize recovery and prevent persistent pain and disability. It directly targets catastrophic thinking about pain (eg, misconceptions about pain and activity, hopelessness, helplessness, and magnifications of pain) and pain anxiety (cognitive, physiological, and pain avoidance elements) by teaching relaxation response and mindfulness skills, correction of misconceptions (eg, through education and adaptive thinking techniques), acceptance and value-based engagement in activities, and activity pacing. TOR introduces and emphasizes skills through didactics, in-session activities, discussions, and home practice assignments. The latest iteration of TOR (following the pilot RCT) consists of four sessions (4 weeks; 45-minute sessions). Home practice assignments will involve practicing TOR skills. Practicing will be facilitated through a web platform that will include all program skills and instructions as individual recordings. The study staff will download this web platform on TOR participants’ phones as a web-based app after randomization. Table 1 shows all the TOR components. For more details on TOR and its components, see the report of pilot feasibility RCT in the paper by Vranceanu et al [26].

The study clinicians are 2 doctoral-level clinical psychologists based at site A, with background and clinical training in mind-body interventions and heterogeneous pain. Clinicians will deliver TOR remotely to the participants at all sites. Makeup sessions with study clinicians will be scheduled for participants who miss any session.
Table 1. Descriptions of the four sessions in the Toolkit for Optimal Recovery, an individual live video–based active intervention.

<table>
<thead>
<tr>
<th>Session</th>
<th>Topics</th>
<th>Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>* Discuss treatment rationale and goals&lt;br&gt;  * Review and correct misconceptions about recovery trajectory after orthopedic injury&lt;br&gt;  * Normalize pain after an injury, move patients away from the mind-body dichotomy by discussing how all pain sensations originate in the brain, and discuss the difference between hurt versus harm&lt;br&gt;  * Learn how the sympathetic nervous system influences symptoms; learn about the disability spiral and how it can lead to slower recovery and chronic pain after orthopedic injury; and learn about the physical, emotional, and cognitive factors that can speed or slow recovery after orthopedic injury&lt;br&gt;  * Provide education about the parasympathetic nervous system and relaxation and demonstrate relaxation strategies (diaphragmatic breathing and body scan)&lt;br&gt;  * Set goals for skills practice: practice 1 relaxation strategy daily</td>
<td>* Pain and recovery misconceptions&lt;br&gt;  * Relaxation response (diaphragmatic breathing and body scan)</td>
</tr>
<tr>
<td>2</td>
<td>* Practice diaphragmatic breathing, review previous material and homework, and problem-solve barriers to practice&lt;br&gt;  * Conduct mindfulness exercise on pain sensations; assist patients in identifying what thoughts, feelings, and behaviors are triggered by the pain sensations and normalize this experience; and provide education about mindfulness techniques for observing thoughts, feelings, and behaviors nonjudgmentally&lt;br&gt;  * Learn decision tree for unhelpful thoughts: adaptive thinking or reframing of thoughts that are not true (eg, “Pain means that I am getting worse”) and acceptance, validation or compassion, and letting go of thoughts previously reframed that keep coming back or those that are true but not helpful (eg, “It is harder to walk right now”)&lt;br&gt;  * Set goals for skills practice: practice diaphragmatic breathing, body scan or mindfulness on pain daily, and complete at least one decision tree exercise</td>
<td>* Mindfulness and meditation&lt;br&gt;  * Adaptive thinking or restructuring versus acceptance of thoughts</td>
</tr>
<tr>
<td>3</td>
<td>* Practice diaphragmatic breathing, review previous material and home practice, and problem-solve barriers to practice&lt;br&gt;  * Provide rationale for activity pacing, assist patients in setting activity goals consistent with their values, and normalize avoidance of activities that are associated with injury and reinforce rationale for approach&lt;br&gt;  * Assist patients in applying acceptance and reframing or problem-solving skills to achieve activity-pacing goals&lt;br&gt;  * Set goals for skills practice: practice diaphragmatic breathing, body scan, or mindfulness daily; complete at least one decision tree exercise, including options for problem solving, acceptance, and reframing; and follow activity-pacing protocol</td>
<td>* Activity pacing&lt;br&gt;  * Value-based living</td>
</tr>
<tr>
<td>4</td>
<td>* Practice diaphragmatic breathing, review previous material and home practice, and problem-solve barriers to practice&lt;br&gt;  * Review all skills, assist patients in identifying which skills are being used, how helpful they are, and how they can be implemented in the future&lt;br&gt;  * Interactive quiz to identify the improvements patient has made, skills that are being used, skills the patient would like to continue to work on, and a plan for continued coping</td>
<td>* Skill consolidation and review</td>
</tr>
</tbody>
</table>

Live Video Delivery

The TOR intervention will be delivered remotely to all participants by the study team at site A. We will use a Health Insurance Portability and Accountability Act–approved secure videoconferencing software (Zoom) routinely used in clinical practice at site A to deliver the intervention. The study team has substantial experience with the implementation of similar mind-body programs through live videos [35-37]. Following randomization, the RA at each site will install the videoconferencing software (Zoom) on the TOR participants’ devices and teach the participants how to use it. We will send a reminder email to the participants for all sessions, including a link to connect to the live video session. An RA will be available to provide real-time support with any connection issues or other technical difficulties that may occur during the intervention sessions.

Treatment Fidelity and Patients’ Safety Considerations

We will adhere to the National Institutes of Health guidelines [38] to ensure treatment fidelity. We will use a structured manual, already developed and used in our previous single-site feasibility trial. The principal investigator (PI; AMV) will train all study clinicians on the study protocol and manualize the TOR intervention. Study clinicians will complete fidelity checklists after each session and receive weekly supervision from the senior psychologist. All study sessions will be audio recorded, and 20% of the audio recordings will be randomly selected to review for fidelity. A total of 2 independent raters will assess the selected sessions against a fidelity checklist to evaluate the fidelity of intervention delivery, with attention to rater discrepancies.
Considerations for Participant Safety During a Virtually Delivered Program

Special safety considerations must be taken to ensure participant safety, accurate data collection, and uninterrupted live video intervention delivery. Our system for virtual program delivery is similar to the current system for site A psychologists and physicians who deliver virtual patient care. As with standard outpatient clinical care at site A, we will have full support from the site A telehealth department to ensure immediate attention to any technical issues. The PI has conducted a live video pilot clinical trial for patients with acute orthopedic trauma and is currently conducting one for patients with neurofibromatosis from across the globe with no technical difficulties. The study team will also receive training in the software, including strategies for resolving potential technical issues. We will set up and test the Zoom software during the baseline visit to the clinic to ensure that technical issues during study visits are minimized. All intervention participants will be instructed on how to rejoin the study visit if the connection is disrupted. The RA will be available to assist clinicians with technology during study visits as needed (eg, helping participants rejoin the session while the clinician continues delivering material). We will closely monitor any severe psychological symptoms, including suicidality, and provide referrals for appropriate levels of care as needed. During enrollment, all intervention patients will be asked to provide names, email addresses, and phone numbers of 2 emergency contact persons for this purpose. We will inform all patients that in the case of suicidality, a warm handoff will be performed to refer patients to local psychological services for a safety evaluation. In the unlikely event of serious concerns of suicidality and to ensure patient safety, we will suspend confidentiality and alert the site PIs so that appropriate clinical interusions can be assured. Safety will always be prioritized over participation in the study.

The MEUC Control

Participants in the MEUC will receive a booklet containing brief, summarized information that reflects the active intervention topics, including the trajectory of pain and recovery after orthopedic illness, the role of relaxation strategies to manage pain, and the importance of returning to engagement in activities of daily living. In addition, similar to participants in the TOR group, participants in this group will receive usual medical care as determined by the medical team. Usual care involves meetings with surgeons, medical staff, pain medications, and physical therapy. Usual care is identical in the intervention and control groups.

Protocol Fidelity and Investigator Compliance

The PI will lead a half-day session with the teams at each of the 3 sites to review the study protocol. The PI will also lead weekly research team meetings with all sites to review study operations and compliance (eg, screening, recruitment, data collection, and management), discuss any emerging concerns, identify training needs, and address any human subject issues. At the end of each study year, the designated Data and Safety Monitoring Committee will audit cases of 5 randomly selected participants and review compliance with study activities at all sites as per the study protocol.

Assessments

We selected the measures and assessment domains in line with the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials and Outcome Measures in Rheumatology guidelines [39], study aims, and guidelines for feasibility trials [40]. All assessments presented below except those indicated otherwise will be collected at baseline (preintervention), postintervention, and at a follow-up timepoint of 3 months.

Sociodemographic Information

We will collect data on age, gender, race, ethnicity, education, employment, income, and marital status using a demographic questionnaire. We will also collect information on mental health history, current psychotropic or pain medication intake, substance use history and status, comorbid medical conditions, and history of depression or other mental health conditions. We will also collect data on smoking status, alcohol use, marijuana intake, and narcotic and nonnarcotic analgesic intake. These assessments will only be conducted at baseline.

Clinical Variables

We will collect data on the injury type, date and location, and injury severity (Abbreviated Injury Severity Index) [41] as rated by the surgeon. These assessments will only be conducted at baseline.

Primary Outcomes

Feasibility of Recruitment

We will calculate the percentage of participants who agree to participate from those who are approached. We will also calculate the subgroup of minorities in each group.

Appropriateness of Treatment

The Credibility and Expectancy Questionnaire [42] assesses treatment expectancy and credibility in clinical outcome studies. This will be completed by participants randomized to the TOR after the delivery of the first session of the intervention.

Feasibility of Data Collection (Self-Report)

We will calculate the percentage of participants who complete the measures at the 3 timepoints (baseline, posttreatment, and 3-month follow-up).

Acceptability of TOR (Adherence to Sessions)

We will report the percentage of participants who are randomized within 1 arm and complete the posttest.

Treatment Satisfaction

The Client Satisfaction Scale [43] assesses participant satisfaction with TOR at posttreatment.

Secondary Outcomes

Feasibility of Randomization or Adherence to the Assigned Arm

We will report the percentage of enrolled participants who start within 1 arm and complete posttest assessments.

https://www.researchprotocols.org/2021/4/e28155
**Fidelity to Study Procedures**
We will report the number of protocol deviations observed per site among all the enrolled participants.

**Adherence to TOR Homework**
We will calculate the ratio of homework logs returned and the number of days in the last week homework was practiced among patients randomized to the TOR (daily self-report log).

**Acceptability as Rated by Therapist**
We will report the percentage of participants who score over 7 on the therapist rating of participation quality in each intervention session among patients randomized to the TOR.

**Feasibility of Data Collection for Performance-Based Measures of Physical Function**
We will report the percentage of enrolled participants who complete performance-based measures of physical function (walk or grip test) at each of the 3 timepoints.

**Feasibility of Data Collection for Rescue Analgesics (Nonnarcotic)**
We will calculate the percentage of enrolled participants with complete data on nonnarcotic rescue analgesics.

**Feasibility of Data Collection for Rescue Analgesics (Narcotic)**
We will calculate the percentage of all participants with complete data on narcotic rescue analgesics.

**Feasibility of Data Collection for Adverse Events**
We will calculate the percentage of enrolled participants with complete data on adverse events (see information on adverse events data collection under the Adverse Events section).

**Therapist-Rated Adherence or Fidelity of the Participants to Session**
We will calculate the percentage of participants who present ≥75% adherence, indexed by the study clinician using checklists and audio recordings at each intervention session.

**Pain**

**The Numerical Rating Scale**
The Numerical Rating Scale (NRS) is a commonly used measure of pain intensity that has demonstrated good psychometric properties [44]. The item assesses pain using an 11-point scale ranging from no pain (0) to worst pain imaginable (10). This measure will be used to assess pain during rest and activity. We will report the sensitivity to detect changes in pain as determined by the percent change in NRS score.

**Within-Group Change in Rescue Analgesics (Narcotic and Nonnarcotic)**
We will calculate the changes in narcotic and nonnarcotic analgesics and concomitant pain treatment (daily self-report log).

**Self-Reported Physical Function**

**Patient-Reported Outcomes Measurement Information System–Physical Function (Version 1.0)**
We will calculate sensitivity to detect changes in self-reported physical function as determined by percent change. This questionnaire assesses one’s ability to perform activities that require physical actions, ranging from low-impact tasks (e.g., self-care, bathing, and dressing) to vigorous physical activities (e.g., running, strenuous sports). The questionnaire consists of 121 Likert scale items, with response options ranging from 1 (without any difficulty or not at all) to 5 (unable to do). The items do not refer to a particular recall period but rather involve the participant’s status at the time of completion [45]. The resulting t score is a standardized score with a mean of 50 (SD 10). Lower scores are indicative of greater disability.

**The Short Musculoskeletal Function Assessment Questionnaire**
We will report sensitivity to detect changes in self-reported physical function as determined by percent change. This questionnaire is a validated 46-item survey that measures physical functioning and musculoskeletal disability [46]. It was developed from a 101-item parent questionnaire with excellent psychometric properties [47]. The individual items are rated on a 4-point Likert scale, with high scores indicative of higher disability. It consists of 2 subscales calculated by summing up the individual items: (1) assessment of function (34 questions) and (2) perception of bothersomeness of the symptoms (12 questions). These sum scores will be transformed to final scores ranging from 0 to 100.

**Performance-Based Physical Function (Walk or Grip Test)**
We will calculate the sensitivity to detect changes in performance-based physical function by percent change. For those with lower extremity or body injuries, we will conduct a timed 10-m walk test [48]. This test is a valid assessment of walking speed when used as a time indicator of health status. The participants will be asked to walk without assistance for 10 m. The time will be measured for the intermediate 6 m, allowing for acceleration and deceleration. Participants will be permitted to use assistive devices as long as they are kept consistent and documented. For those with upper extremity or body injuries, a grip strength test [49] will be conducted. This test is a commonly used simple measure of muscle strength. The strength of the participants’ grip will be measured quantitatively using a hand dynamometer [49].

**Coping and Emotional Function**
We calculated sensitivity to detect changes in coping and emotional function, as determined by the percent change in these outcomes using the following measures.

**The Pain Catastrophizing Scale**
The Pain Catastrophizing Scale (PCS) is a measure of negative pain-related cognitions or catastrophic thinking (exaggerated reporting and rumination about pain), which has demonstrated good psychometric properties [33]. The PCS includes 13 Likert scale items with 4 points ranging from not at all (0) to all the time (3). The PCS consists of 3 subscales: (1) rumination (ie,
repetitively going over thoughts and feelings about pain), (2) helplessness (ie, subjective feeling of hopelessness or helplessness related to the experience of pain), and (3) magnification (ie, exaggerated thinking about the negative consequences of pain). A total score will be computed, with higher scores reflecting poorer coping ability. A score ≥20 on PCS has been proposed as an indicator of high psychosocial risk for poor outcomes after orthopedic trauma and will be used as one of the inclusion criteria in this trial [50]. PCS will be administered before every assessment and intervention session.

The Pain Anxiety Symptom Scale
The Pain Anxiety Symptom Scale (PASS-20) [34] is a measure of fear and anxiety related to the experience of pain. PASS-20 consists of 20 items rated on a 6-point Likert scale from never (0) to always (5). The PASS-20 consists of 4 subscales: (1) avoidance (ie, avoiding activities that cause pain), (2) fearful thinking (ie, fear thoughts related to pain), (3) cognitive anxiety (ie, compromised cognitive function when in pain), and (4) physiological response (ie, somatic anxiety symptoms in response to pain). A total score will also be computed, with higher scores reflecting poorer coping ability. A score ≥40 on the PASS-20 is proposed as an indicator of high psychosocial risk for poor outcomes after orthopedic trauma and will be used as an inclusion criterion in this trial [51,52]. PASS-20 will be administered before every assessment and intervention session.

Measures of Current Status
The Measures of Current Status (MOCS) [53] assesses the perceived ability to use healthy coping skills. The MOCS consists of 4 subscales: (1) relaxation, (2) awareness of stress, (3) assertiveness, and (4) disputing maladaptive thoughts. The MOCS consists of 13 items rated on a 5-point Likert scale ranging from 0 (I cannot do this at all) to 4 (I can do this extremely well). A total score will be computed with scores ranging from 0 to 52, with higher scores reflecting more effective coping ability.

The Posttraumatic Stress Disorder Checklist: Civilian Version
The Posttraumatic Stress Disorder Checklist (PCL) is a measure of posttraumatic stress symptoms with good psychometric properties [54]. The PCL consists of 17 items rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (all the time). The items measure the extent to which the participant has been bothered by symptoms of posttraumatic stress in the past month. The measure provides a total severity score as well as a diagnostic cutoff.

The Center for Epidemiologic Study of Depression
The Center for Epidemiologic Study of Depression [55] is a commonly used measure of depressive symptoms. It consists of 20 items rated on a 4-point Likert scale from rarely or none of the time (less than 1 day) (0) to most or all the time (5-7 days). The measure provides a total severity score for depressive symptoms.

Adverse Events
We will calculate the number of adverse events among all enrolled participants (data collected at each assessment and intervention timepoints).

Feasibility of Collecting Orthopedic Staff Satisfaction Measures
We will calculate the percentage of surgeons and study staff who complete the measures related to staff satisfaction with the study procedures.

Feasibility of Obtaining Data on Orthopedic Staff Perceived Ease of Referrals
We will report the percentage of surgeons and study staff who complete the measure.

Feasibility of Obtaining Data on Orthopedic Staff Perceived Cost-Benefit
We will report the percentage of surgeons and study staff who complete the measure.

Orthopedic Staff Feasibility of Referral
We will calculate the percentage of surgeons who make at least 5 referrals per site.

Feasibility of Obtaining Data on Feasibility of Study Implementation From Staff
We will calculate the percentage of surgeons and staff who complete the data.

Feasibility of Collecting Appropriateness Measures by Staff
We will report the percentage of surgeons and staff who complete appropriateness measure.

Perceived Acceptability of Study by Staff
We will report the percentage of surgeons and staff who score over 7 on a 10-item Likert scale satisfaction measure, among those who completed the measure (collected at the end of the study).

Perceived Acceptability of Staff Regarding Ease of Referral
We will calculate the percentage of surgeons and staff who score over 7 on a 10-item Likert scale ease of referral measure, among those who completed the measure (collected at the end of the study).

Perceived Acceptability of Staff Regarding Cost-Benefit
We will calculate the percentage of surgeons and staff who score over 7 on a 10-item Likert scale cost-benefit measure, among those who completed the measure (collected at the end of the study).

Feasibility of Study Implementation as Perceived by Study Staff
We will calculate the percentage of surgeons and staff who score over 7 on a 10-item Likert scale of feasibility of study implementation as perceived by study staff, among those who completed the measure (collected at the end of the study).

 Appropriateness as Perceived by Study Staff
We will calculate the percentage of surgeons and staff who score over 7 on a 10-item Likert scale of appropriateness, among...
those who completed the measure (collected at the end of the study).

**Analysis Plan and Sample Size Considerations**

We will evaluate the feasibility of each site based on the number of patients identified via medical records, the number referred by the surgeon, number screened, number consented, and number enrolled and randomized. The results will be reported using descriptive statistics (ie, numbers and proportions with 95% CI).

The target sample size of 60 participants enrolled per site (180 participants in total) will provide sufficient data to evaluate the feasibility benchmarks required for future hybrid efficacy-effectiveness studies. With 1:1 randomization, we will have 30 participants per arm, which is the minimum required to adequately evaluate measures of fidelity, acceptability, appropriateness, and feasibility. This trial is neither powered for efficacy nor aimed to provide such information. Consistent with the feasibility design of this trial, we will report the means and SDs of all measures at all timepoints, including the distribution of scores and internal consistency. To determine the measures’ sensitivity to detect changes, we will report the percent change in all quantitative outcomes. We will also summarize demographic and clinical variables. Although we will not conduct any efficacy analyses consistent with previous research recommendations [40], we will use estimates of person-to-person variation in efficacy outcomes to determine the required sample size for a future hybrid efficacy-effectiveness trial.

The following benchmarks will be required to be met to proceed to implement a fully powered RCT:

- Overall, ≥70% of patients approached will have to agree to participate (feasibility of recruitment).
- Among all surgeons and study staff: ≥70% of staff will provide data on the feasibility and appropriateness of study implementation, staff satisfaction with study procedures, perceived ease of referral, and cost-benefit. In addition, to achieve feasibility of referral, ≥80% of surgeons need to make at least 5 referrals per site.
- Among surgeons and study staff who complete measures: ≥70% of staff will endorse a score of over 7 on acceptability, defined as staff satisfaction with study procedures, perceived ease of referral and cost-benefit, and feasibility and appropriateness of study implementation.
- Among all patients with orthopedic injuries: ≥70% who were approached agreed to participate, with <5 protocol deviations per site.
- Among all enrolled patients with orthopedic injuries: ≥70% of participants will have to complete self-report data at each of the 3 timepoints; ≥70% of participants who start within one arm must complete posttest; ≥70% of patients must agree to complete walk tests and grip tests; and ≥70% of participants will provide data on narcotic and nonnarcotic analgesics and adverse events, and we will observe stability of nonnarcotic analgesics, decrease in narcotic analgesics, and minimal adverse effect during the study period.
- Among patients randomized to TOR: ≥70% of participants will endorse a score over the scale’s midpoint on credibility and expectancy score and client satisfaction score, ≥3 out of 4 sessions attended by ≥70% of participants, ≥1 out of 3 homework logs returned with 4 out of 7 days of practice by ≥70% of participants, and ≥75% of sessions with a score over 7 on the therapist rating of participation quality.
- We will observe ≥75% adherence by the therapist (indexed through checklist and audio recordings) for ≥70% of the participants.

Once these data analyses are completed, the multidisciplinary team will review the data and discuss the interpretation of our findings in the context of current research on acute orthopedic injury-related pain and disability. In the event that a given benchmark is not met based on the preliminary data collected, site-specific modifications will be made.

**Data Management**

The research team will develop standard operating procedures for data collection, data management, and quality control. The PI (along with 2 co-PIs) will oversee all procedures related to secure data collection and management. Data activities will be conducted using Research Electronic Data Capture (REDCap), a software and workflow tool for research data collection and management, with assistance from the Partners HealthCare Research Computing, Enterprise Research Infrastructure and Services group. REDCap provides a secure web-based app with intuitive user interface data and real-time validation rules. All data will be stored in REDCap at site A.

**Data Quality**

Per study year, the data safety and monitoring committee will receive a data quality report for each site and per randomized group, including a summary of (1) adverse events per randomized group, including the number and types of events, severity, timepoint, and determination of whether the events were study related; (2) an unblinded report of study retention and reasons for attrition per randomized group; and (3) a data quality report for each site, including a summary of participant enrollment and retention, participant adherence with visits, and data completeness and quality. The committee will use this report to evaluate the study’s capacity for valid analysis.

**RCT Deliverables**

We will report on feasibility outcomes for TOR delivery at each site, including the ability to recruit diverse racial, ethnic, and low socioeconomic status populations. We will also obtain the required sample size for a future hybrid efficacy-effectiveness trial.

**Results**

We have initiated the pre-enrollment procedures for the feasibility RCT at each site, with enrollment set to begin in June 2021. The target date of completion for the feasibility RCT is April 2024.
Discussion

Principal Findings

This paper describes the study design and specific strategies used to conduct a multisite feasibility RCT of TOR administered via live videoconferencing to adult patients with acute orthopedic injuries across 3 sites. We provide details on assessing the feasibility of the refined program at 3 heterogeneous level 1 trauma centers that involve demographically diverse patients and implement different guidelines and routines, including overcoming typical barriers to psychosocial care by delivering the intervention via a secure live video. This information is invaluable for future research and provides a novel paradigm for the delivery of preventive care to patients with acute orthopedic injuries.

The long-term goal of this project is to set the stage for a multisite hybrid efficacy-effectiveness trial of TOR versus MEUC. Informed and refined by the findings of the focus groups conducted among orthopedic medical provider stakeholders, the results of this study will address the challenges concerning barriers to biopsychosocial care for orthopedic patients using an implementation framework. Thus, in addition to the use of novel methods such as delivering the intervention via secure live video, we also revised the initially proposed study protocol (mainly recruitment strategy and inclusion and exclusion criteria) and developed educational materials for the care team based on focus group data to maximize success at each site. We will also be able to address some limitations of the single-site feasibility RCT, such as the lack of ethnic, racial, and socioeconomic diversity in the study population. This line of work has the potential for adaptation and generalizability to surgical orthopedic centers and even primary care practices.

The findings will also provide a realistic examination of the study methods, as they will take place in a fully powered, hybrid efficacy-effectiveness trial. Specifically, by seeking feedback from providers and staff during the commencement of this study, we have gained knowledge regarding the potential site-specific differences that need to be addressed for future hybrid efficacy-effectiveness trials and dissemination of our effective intervention. Our efforts to refine the TOR program based on this feedback and achieve a priori feasibility benchmarks of this intervention. Our guiding hypotheses for this next trial are that receiving TOR delivered over live video at 1 to 2 months after an orthopedic injury will be associated with improvements in pain and disability, catastrophic thinking about pain, pain anxiety, and depressive and posttraumatic symptoms, compared with MEUC. Using the findings of this hybrid efficacy-effectiveness trial, we aim to set the stage for dissemination studies.

Foreseen Challenges

Despite the innovative approach of this project, some potential challenges warrant consideration. First, the recruitment of ethnically diverse samples could be challenging. The 3 sites of this study are geographically dispersed and encompass a diverse sociodemographic sample. We will also approach every ethnic and racial minority participant and prioritize their recruitment (see the Methods section). If by the end of year 2, we do not end up with at least 25% minorities in our study samples across sites, we will expand our recruitment to a preidentified urban medical center with 80% racial and ethnic minority patients. Second, there are concerns regarding the competing research projects. To overcome this barrier, we have secured support from orthopedic leaders at each site to prioritize this study at each site. Third, special considerations exist regarding live video interventions in terms of participants’ safety, accurate data collection, and uninterrupted delivery of the live video intervention. In addition to having the full support of site A telehealth, we have ensured that our study team has adequate training in patient safety considerations, use of the software, and problem solving of any potential technical issues.

Implications

The recognition of the pivotal role of psychosocial interventions in orthopedic care is increasing. However, this awareness has yet to be reflected in the development of feasible psychosocial interventions to address this need or in real-life practices of the surgeons and related medical staff who do not reliably refer their patients to receive psychosocial support. This multisite feasibility RCT is a part of the work toward developing the first preventive psychosocial intervention to address the development of chronic pain and disability among at-risk patients with acute orthopedic injuries. The results will inform the fully powered hybrid efficacy-effectiveness trial, which will utilize the benchmarks and program refinements established by this study. The goal is to improve the care of orthopedic patients with musculoskeletal injuries who are at risk for chronic pain and disability, by shifting the paradigm of care for these patients from a purely medical model to a biopsychosocial model with provision of psychosocial interventions that are acceptable (by the patients and care team) and easily delivered within all orthopedic trauma practices. Such a paradigm shift will provide a new way to improve patient and provider satisfaction, decrease suffering, and decrease health care costs among this group of patients. Moreover, if effective, this model of care and related implementation framework can be applied in other medical settings.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Peer-review report by the National Center for Complementary and Integrative Health Special Emphasis Panel, Exploratory Clinical Trials of Mind and Body Interventions (National Institutes of Health).
References


35. Vranceanu et al JMIR RESEARCH PROTOCOLS


53. Carver C. Measure of Current Status Internet. College of Arts and Sciences: Psychology. URL: https://local.psy.miami.edu/faculty/ccarver/sclMOCS.phtml


**Abbreviations**

IRB: institutional review board
MEUC: minimally enhanced usual care
MOCS: Measures of Current Status
NRS: Numerical Rating Scale
PASS-20: Pain Anxiety Symptom Scale
PCL: Posttraumatic Stress Disorder Checklist
PCS: Pain Catastrophizing Scale
PI: principal investigator
RA: research assistant
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
TOR: Toolkit for Optimal Recovery

Edited by T Derrick; submitted 23.02.21; this is a non–peer-reviewed article; accepted 09.03.21; published 28.04.21.

Please cite as:
Vranceanu AM, Bakhshaie J, Reichman M, Doorley J, Elvey AR, Jacobs C, Chen N, Esposito J, Laverty D, Matuszewski PE, Fatehi A, Bowers LC, Harris M, Ring D
A Live Video Program to Prevent Chronic Pain and Disability in At-Risk Adults With Acute Orthopedic Injuries (Toolkit for Optimal Recovery): Protocol for a Multisite Feasibility Study
JMIR Res Protoc 2021;10(4):e28155
URL: https://www.researchprotocols.org/2021/4/e28155
doi:10.2196/28155
PMID:33908886
Protocol

Development of a Novel Intervention (Mindful Steps) to Promote Long-Term Walking Behavior in Chronic Cardiopulmonary Disease: Protocol for a Randomized Controlled Trial

Daniel Litrownik1, BA; Elizabeth A Gilliam1, MA; Peter M Wayne2, PhD; Caroline R Richardson3, MD; Reema Kadri3, MLIS; Pamela M Rist2, ScD; Marilyn L Moy4, MD, MSc; Gloria Y Yeh1,2, MD, MPH

1Division of General Medicine, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, United States
2Osher Center for Integrative Medicine, Harvard Medical School and Brigham and Women’s Hospital, Boston, MA, United States
3Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States
4Pulmonary and Critical Care Section, Department of Medicine, Veterans Administration Boston Healthcare System, Boston, MA, United States

Corresponding Author:
Daniel Litrownik, BA
Division of General Medicine
Department of Medicine
Beth Israel Deaconess Medical Center
330 Brookline Ave
Boston, MA
United States
Phone: 1 617 754 1423
Email: DLIITROW@gmail.com

Abstract

Background: Despite current rehabilitation programs, long-term engagement in physical activity remains a significant challenge for patients with chronic obstructive pulmonary disease (COPD) and heart failure (HF). Novel strategies to promote physical activity in these populations are greatly needed. Emerging literature on the benefits of both mind–body interventions and web-based interventions provide the rationale for the development of the Mindful Steps intervention for increasing walking behavior.

Objective: This study aims to develop a novel multimodal mind–body exercise intervention through adaptation of an existing web-based physical activity intervention and incorporation of mind–body exercise, and to pilot test the delivery of the new intervention, Mindful Steps, in a randomized controlled feasibility trial in older adults with COPD and/or HF.

Methods: In phase 1, guided by a theoretical conceptual model and review of the literature on facilitators and barriers of physical activity in COPD and HF, we convened an expert panel of researchers, mind–body practitioners, and clinicians to inform development of the novel, multimodal intervention. In phase 2, we are conducting a pilot randomized controlled feasibility trial of the Mindful Steps intervention that includes in-person mind–body exercise classes, an educational website, online mind–body videos, and a pedometer with step-count feedback and goals to increase walking behavior in patients with COPD and/or HF. Outcomes include feasibility measures as well as patient-centered measures.

Results: The study is currently ongoing. Phase 1 intervention development was completed in March 2019, and phase 2 data collection began in April 2019.

Conclusions: Through the integration of components from a web-based physical activity intervention and mind–body exercise, we created a novel, multimodal program to impact long-term physical activity engagement for individuals with COPD and HF. This developmental work and pilot study will provide valuable information needed to design a future clinical trial assessing efficacy of this multimodal approach.

Trial Registration: ClinicalTrials.gov NCT03003780; https://clinicaltrials.gov/ct2/show/NCT03003780

International Registered Report Identifier (IRRID): DERR1-10.2196/27826

(JMIR Res Protoc 2021;10(4):e27826) doi:10.2196/27826

KEYWORDS
mind–body exercise; internet-mediated intervention; behavior change; physical activity; COPD; heart failure

https://www.researchprotocols.org/2021/4/e27826
Introduction

Physical activity is an important modifiable behavior that has enormous impacts on cardiopulmonary health. Chronic obstructive pulmonary disease (COPD) and heart failure (HF) are two systemic cardiopulmonary syndromes where patients experience similar morbidity and suffer debilitating decreases in physical activity. COPD, characterized by progressive airflow limitation most commonly due to environmental toxins (eg, smoking), confers considerable morbidity in up to 10% of US adults and is the third most common cause of death in the United States [1,2]. HF, also a progressive syndrome, is characterized by the inability of the heart to pump efficiently to meet metabolic demands and is associated with multiple cardiovascular and metabolic derangements [3,4]. Coexistence of the 2 conditions is being increasingly recognized, with estimated rates as high as 39% [5,6].

Patients with COPD and HF are characteristically deconditioned, with impairments in exercise tolerance and increases in dyspnea. Both self-reported and directly measured physical activity is significantly reduced, even at the earliest stages of disease [6-10]. During and following acute exacerbations, patients suffer further dyspnea, deconditioning, and reductions in physical activity [11-13]. Both cardiopulmonary conditions are also associated with multiple comorbidities, including anxiety, depression, and musculoskeletal disease, further contributing to reduced physical activity [6,14-16]. In HF, systematic reviews and meta-analyses report that exercise training reduces HF-related hospitalizations and results in clinically important improvements in health-related quality of life (HRQL) [17,18]. Walking, the most common form of exercise in these populations, has been associated with reduced HF risk [19]. The least active individuals with COPD have risks that are 2-6 times higher for acute exacerbations and hospitalizations than those most active [20]. Higher physical activity in COPD is also associated with a significantly lower hospital readmission rate and mortality risk, independent of lung function [21-23]. COPD studies examining daily step count support that every step walked can positively impact disease course [20,21,24,25].

Unfortunately, engagement in physical activity is a universal challenge amplified in these chronic cardiopulmonary populations. Conventional center-based cardiac and pulmonary rehabilitation have been shown to improve outcomes; however, they are vastly underutilized, not accessible to the population at large, and not sustainable [1-6,16,26-28]. Providers refer less than 13% of potential candidates who would benefit from pulmonary or cardiac rehabilitation [29-32], and among those who are referred, noncompletion rates are as high as 20%-40% [31,33-36]. Despite improvements in exercise capacity, dyspnea, and HRQL after a typical course, benefits diminish 6-12 months after program completion [37]. Other outpatient and home-based programs have had variable success [38-40]. Qualitative studies have identified barriers and facilitators to long-term adherence to physical activity in COPD and HF. Important themes include addressing fears, promoting confidence, implementing personalized feedback and goals, providing peer support, and fostering a conducive environment with opportunities to engage in exercise [41,42]. Walking, in particular, has been identified as simple and accessible. Within this context, there has been interest in refocusing the current paradigm from one of promoting short-term aerobic exercise in structured settings to one of sustaining long-term everyday physical activity. Novel strategies to achieve this in these cardiopulmonary populations are greatly needed.

There is a burgeoning interest and emerging literature in mind–body interventions for fostering positive behavior change and physical activity [43-45]. In chronic cardiopulmonary disease, mind–body exercise, such as tai chi and yoga, may be particularly relevant [46,47]. Research has suggested that tai chi and mind–body breathing exercises are safe and feasible in older deconditioned patients and can lead to increases in HRQL, mood, exercise self-efficacy, and overall physical activity. In studies specific to COPD and HF, self-efficacy and overall empowerment have been shown to be important factors in facilitating long-term behavior change [40,48].

In addition, there is a rapidly emerging literature on use of web-based technology to promote healthy habits and behavior change [49-69]. Some home-based programs and lifestyle physical activity interventions, which combine supervised and independent exercise with self-monitoring devices, such as pedometers, have shown success in increasing physical activity [70-73]. One such web-based intervention, made specifically for elderly patients with COPD, increased HRQL and physical activity (daily step counts) in the short term (4-6 months) [74,75]. This multimodal intervention (Every Step Counts) created by Moy and Richardson includes an online interactive web platform coupled with a pedometer, step goal feedback, motivational and educational content addressing barriers to exercise, and an online community forum for social support. Development of this intervention was based on the behavioral theory of self-regulation, which emphasizes an iterative process of behavior change with individualized goal setting, iterative feedback, extrinsic motivating factors, and social support [76-78]. However, in long-term follow-up (12 months), benefits from this intervention were not sustained [79].

We hypothesized that by modifying this intervention with the addition and integration of mind–body principles, we could enhance existing key behavioral constructs and target new constructs through self-reflection, awareness, and personal transformation, which might lead to enhanced self-efficacy and the ultimate outcome of long-term adherence to walking behavior for COPD and HF. Thus, the current developmental study includes 2 phases with the following overarching aims: (1) adapt/refine the existing web-based physical activity intervention with added emphasis on mind–body principles and couple this platform with in-person mind–body exercise to create Mindful Steps; and (2) test delivery of Mindful Steps, a novel multimodal mind–body exercise intervention in a randomized feasibility trial with usual care control (N=42) in patients with COPD and/or HF.
Methods

Phase 1: Intervention Development

Conceptual Framework

We chose self-efficacy as a key outcome of interest. According to social cognitive theory, self-efficacy is one’s confidence in the ability to perform certain health behaviors and can influence the engagement in and actual performance of those behaviors, affecting health outcomes thereby [80]. Self-efficacy is a modifiable characteristic, and exercise self-efficacy is one of the strongest predictors of physical activity behavior [81]. Within our conceptual model, there are multiple characteristics of mind–body exercise that reinforce and enhance constructs of self-regulation from the existing intervention and additionally provide new constructs and potential mediating factors toward self-efficacy [82-84].

Experience/goal attainment, based on traditional self-efficacy theory [80], refers to the notion that prior successes boost confidence and facilitate self-efficacy. For example, mind–body exercise is often described as gentle, accessible, and nonthreatening, which may allow participants to achieve goals, even if in small increments. Similarly, small individual successes in iterative step-count goals further enhance that model with repeated success. Modeling describes when participants gain from seeing others succeed and modeling others’ behaviors. Prior qualitative data, including our studies of tai chi and HF, have reinforced the value of seeing others, perhaps even in a poorer condition than oneself, achieve success, which feeds confidence and motivation for one’s own abilities [82,83]. Self-awareness and mindfulness of physiologic cues and emotional states are inherently cultivated through mind–body exercise and are at the foundation of mind–body practices. Body signals such as fast heart rate or breathing due to anxiety or fear can negatively impact one’s perception of a particular task, such as exercise or walking. This may be illustrated by the anxiety–breathlessness cycle in COPD [85]. Strategies for mindfulness, body– and self-awareness, and concomitant acceptance and nonreactivity can mitigate this effect. Mind–body exercises can also address pain cycles and related exercise barriers [86-89] and impact fear of falling, which is correlated with self-efficacy and decreased falls [90,91]. Intrinsic motivation refers to motivation that is driven, not because of external factors (eg, a reward through an intervention website, feedback, or praise) but by internal factors, such as personal satisfaction and enjoyment. The inherent likeability of mind–body therapies is not well studied; however, numerous studies have reported qualitatively on enjoyment of mind–body exercise [82,84,88,92]. Certainly, increased mood with physical activity is well established. Finally, each of these potentially mediating factors are held within a larger domain of self-reflection and transformation that is a part of self-determination theory [93]. There is a growing literature that describes personal transformative experiences that can occur with mind–body practices [44–46].

In the modified Mindful Steps intervention, we retain the original constructs of goal setting, iterative feedback, extrinsic motivation, and social support that have appeared effective for short-term adherence and additionally introduce constructs that facilitate self-reflection and personal transformation through heightened, mindful, self-awareness and intrinsic motivation that can impact overall self-efficacy and facilitate long-term physical activity behavior [80,93,94].

Expert Panel

We assembled an expert panel to inform the creation and integration of a mind–body curriculum within the existing web-based platform. The panel included those with clinical and research expertise in mind–body exercise and tai chi, chronic cardiopulmonary disease (COPD and HF specialists), behavioral psychology, exercise physiology, and physical activity. We first...
began with open feedback from the core investigator team to develop the initial proposal for the curriculum. Using a modified Delphi process [95,96], we then solicited input from the larger expert panel on issues from practical implementation (eg, optimum ratio of and integration of in-person class contact with online video to decrease participant burden, modifications for safety, use of technology) to content-specific recommendations (eg, essential components of the mind–body curriculum, how best to integrate it with existing content, modifying disease-specific educational content for HF and already existing information for COPD). After a first round expert panel group meeting, individual interviews took place in an iterative fashion to gather further input and refine the protocol. After 2 subsequent group meetings with the expert panel, we arrived at a consensus protocol outlined in the following sections.

**Mindful Steps: Intervention Components**

The final Mindful Steps Intervention to be pilot-tested contains 7 primary components (Table 1). Each of the components is described below, with specific information on how it was modified from the prior Moy and Richardson intervention.
<table>
<thead>
<tr>
<th>Intervention component</th>
<th>Description</th>
<th>Modifications from prior intervention</th>
<th>Rationale</th>
<th>Theoretical constructs</th>
</tr>
</thead>
</table>
| 1. Pedometer–website interface with individualized step goal | ● Fitbit HR graphical display of daily/weekly steps  
● Individualized daily step goals based on prior weekly step counts | ● No change                           | ● Allows accurate self-monitoring and personalized feedback of real-time progress  
● Provides motivation to increase walking behavior  
● Small, frequent successes in goal attainment leading to future successes | ● Goal setting  
● Experience/goal attainment  
● Extrinsic motivation  
● Iterative feedback  
● Self-efficacy |
| 2. Motivational/educational website content | ● Motivational messages promoting walking and overcoming barriers  
● Educational tips on disease self-management | ● Relevance of content expanded to HF population (in addition to COPD)  
● Integrated mind–body content  
● Expanded second 6 months with motivational memes and links to mind–body videos | ● Provides practical tools and strategies to promote walking, manage illness, and overcome barriers to walking | ● Extrinsic motivation  
● Intrinsic motivation  
● Self-efficacy  
● Self-awareness and mindfulness |
| 3. Online forum | ● Private online forum for study participants | ● Use of the forum by intervention instructors to facilitate engagement, sharing, and encourage walking | ● Provides community, social support, and sense of continuity with in-person classes | ● Social support  
● Modeling  
● Self-efficacy |
| 4. Mindful walking curriculum videos | ● Weekly short educational video clips (2-8 minutes); theme-based (see section Component 4: Mindful Walking Video Curriculum)  
● Short didactic teaching and guided mind–body exercises | ● New component | ● Provides practical tools and strategies for overcoming barriers to walking  
● Cognitive reframing to find joy in walking  
● Integrates mindfulness practices with everyday walking | ● Self-efficacy  
● Intrinsic motivation  
● Self-awareness and mindfulness |
| 5. Mind–body exercise video library | ● Short videos (5-10 minutes) of mind–body exercises that support walking, including both standing and seated meditative exercises | ● New component | ● Reinforces in-person class learning  
● Cultivates breath and body interoceptive awareness | ● Self-efficacy  
● Self-awareness and mindfulness |
| 6. Live group classes | ● In-person group class following Mindful Walking curriculum (see section Component 4: Mindful Walking Video Curriculum)  
● Incorporating mindful walking, mind–body exercises, and facilitated group discussion | ● New component | ● Supports engagement with the website and step goal progress  
● Allows time to unpack themes of weekly videos, practice and refine mind–body exercises  
● Encourages social modeling of successes  
● Provides community and social support | ● Self-efficacy  
● Intrinsic motivation  
● Self-awareness and mindfulness  
● Modeling  
● Experience/goal attainment  
● Social support |
Theoretical constructs

Rationale

7. Earn your stars

<table>
<thead>
<tr>
<th>Intervention component</th>
<th>Description</th>
<th>Modifications from prior intervention</th>
<th>Rationale</th>
<th>Theoretical constructs</th>
</tr>
</thead>
</table>
| • Extrinsic motivation | - Online reward system that rewards reaching step goals, watching videos, and seeing motivational/educational web content | • New component | • Positive reinforcement of website engagement and walking behavior | • Extrinsic motivation  
• Self-efficacy |
| • Self-efficacy | | | | |
| • Positive reinforcement of website engagement and walking behavior | | | | |
| • New component | | | | |
| • Online reward system that rewards reaching step goals, watching videos, and seeing motivational/educational web content | | | | |
| • Small prizes for reaching 100 star and 500 star milestones | | | | |

Component 1: Pedometer–Website Interface and Individualized Step-Count Goals

The pedometer–website interface and individualized step-count goal components are unchanged from the original Moy and Richardson intervention [97]. Participants each receive a Fitbit Alta HR with a wristband and a personal account with access to the study web platform. An algorithm developed in prior studies is used to calculate individualized daily step-count goals each week, taking the average of the daily step counts from the prior week (most recent 7-day period of valid step-count data) and adding 400 steps [64,65,98,99]. Step-count goals do not exceed 10,000 steps per day. A graphical display of daily and weekly step counts and individual goals are highlighted on the home page allowing subjects to self-monitor progress in real time (Figure 2). Participants are asked to wear the pedometer at all times during waking hours and to sync the device daily for up-to-date feedback. Participants are encouraged to log into the website daily and also receive an automated weekly email with their new step-count goal for the coming week. Step-count goals provide extrinsic motivation to increase walking behavior over time.

Figure 2. Mindful Steps home page.

Component 2: Motivational and Educational Website Content

The Motivational Messages and Educational Tips components are practical intervention tools and strategies to promote walking, manage symptoms of disease, and overcome barriers to physical activity. Motivational Messages (weekly message, total 52) includes titles such as “Incorporating Walking into your Daily Life,” “Managing Stress and Anxiety,” “Handling Set-Backs,” and “Keep Walking Fun.” Educational Tips (one tip every other day, total 88) consists of brief education information and includes titles such as “Use Your Shortness of Breath to Pace Yourself,” “Using Your Body Posture to Help You Feel Less Short of Breath,” “When Is Muscle Soreness Good?,” “Top 10 Reasons to Walk,” and “People Who Walk More Live Longer.” Educational Tips also includes links to external, publicly available websites that provide further material to explore the proposed topic (eg, the Heart Failure Society of America website information, “What if your Breathing Symptoms Worsen,” on exercise and symptom management for HF). Both Motivational Messages and Educational Tips were further adapted to include links to Mindful Walking curriculum videos (component 4) and mind–body exercise
videos (component 5) as appropriate. Additionally, the Motivational Message component was modified to include prompts throughout for mindful awareness, brief guided mindfulness practices, emotion regulation techniques to address anxiety and fear, cognitive reframing of walking from a chore to a gift, and practical information on how and where to walk safely.

Component 3: Online Forum
A private forum was available on the web platform in the original Moy and Richardson intervention, but was underused [79]. In Mindful Steps, we aim to use the forum as an opportunity to engage participants in between live group classes (component 6). Prompts from class instructors and study staff are given to encourage dialogue and reflection on experiences with walking and further explore themes from classes, videos, Motivational Messages, and Educational Tips. The forum also provides an opportunity for participants to ask questions to the larger group, share successes and barriers to their daily walking and mindful practices, and offer helpful practical information, such as their favorite indoor walking spaces or outdoor trails. The purpose of the forum is to provide social support, social modeling of successes, and encourage community.

Component 4: Mindful Walking Video Curriculum
Based on feedback from our expert panel, we created brief (5-10 minute) videos designed to support walking as exercise, with each video centered around a particular theme (listed below).

Theme
1. Introduction: Body, Mind, and Breath
2. Mindful Warm-ups: Lower Body
3. Motivation to Move
4. Putting Joy Back into Exercise
5. Rewarding Yourself with the Gift of Walking
6. Mindfulness in Motion
7. Mindful Warm-ups: Upper Body
8. Renew Your Body with Your Breath
9. Every Step Counts
10. Self-Kindness
11. Overcoming Barriers and Challenges
12. Walking for Your Mind and Spirit
13. Pain Management
14. Preventing Falls
15. Strength and Flexibility
16. Breath Awareness
17. Your Body Affects Your Mind
18. Stop and Smell the Roses
19. Relaxation and Stress Management
20. Exploring Balance in Walking
21. Belly Breathing
22. Rest and Recharge along the Way
23. Importance of Posture
24. Leg Strength
25. Moderated Effort: the 70% Rule
26. Mechanics of Walking

A new video populates the Mindful Steps home page weekly (total 26; videos are repeated in the second 26 weeks). The video includes a short didactic teaching on the theme followed by theme-related guided mind–body exercises. For example, the Relaxation and Stress Management video includes a short teaching on the impact that mindful breath awareness can have on stress, mood, and symptoms of dyspnea, and is followed by a guided mindful breathing practice. The video concludes by encouraging viewers to practice mindful breathing while walking and to integrate these practices into activities of daily living.

The full video curriculum provides practical tools and strategies for overcoming barriers to walking by increasing awareness of cognitive, emotional, and bodily cues, and learning skillful ways to respond to these barriers, including pausing and slowing down, acceptance and nonreactivity, meditative breathing, active relaxation, and gentle stretches that alleviate common aches and pains. One recurring theme (also present in component 2) is the reframing of walking and exercise from being a chore or burden to being a gift, emphasizing the enjoyment and pleasure of walking, and thus cultivating a shift towards intrinsically motivated walking. The videos also support the application of mindfulness and mind–body techniques to walking itself as a way to support a more enjoyable, embodied, walking experience.

Component 5: Mind–Body Exercise Video Library
A library of 13 mind–body exercise videos used in our prior tai chi trials [100-103] are available on the website. These include the following: Swinging and Drumming, Swinging Up and Down, Hip and Leg Circles, Stretching Hands and Wrists, Spinal Cord Breathing, Song Breathing Shoulder Lifts, Dragon Wags Its Tail, The Fountain, Washing Your Body with Healing Energy, Renewing Your Body with Your Breath, Mindful Breathing, Belly Breathing, and Kindness Meditation. These videos reinforce the in-class learning of the mind–body exercises classes (component 6) and facilitate participants’ home practice of mind–body exercises. Collectively, the exercises include gentle aerobics, stretching, and meditative breathing, and aim to increase the range of motion and bring mindful awareness to the breath and individual parts of the body. They emphasize the importance of intention, facilitating ease and enjoyment of motion, and overall relaxation of body and mind.

Component 6: Live Group Classes
Live 75-minute group classes are offered to participants over the course of the year, with more instruction towards the beginning of the intervention period. Classes are conducted weekly for the first 6 weeks, biweekly for the next 34 weeks, and monthly for the last 12 weeks (25 total). Participants attend classes together as a cohort. The class curriculum follows themes from the Mindful Walking curriculum (component 4), including guided mindful walking, and also incorporates mind–body exercises from the video library (component 5). Classes begin and end with facilitated group discussion of class themes, exploration of barriers and successes with regard to step-count goals, and general check-ins. Classes are taught by experienced instructors who have led classes in our prior mind–body exercise clinical trials for over 10 years.

Component 7: Earn Your Stars
Earn Your Stars is an extrinsic motivational reward system to incentivize participants to stay engaged with the website content and their step-count goals. Participants can earn up to 3 stars a
day for each of the following: reaching daily step-count goal, watching a video, clicking on Motivational Messages or Educational Tips. The star counter displays the total star accumulation on the individual’s home page. Small prizes (e.g., US $5 gift cards) are given out when participants reach >100 and >500 stars.

**Phase 2: Pilot Randomized Controlled Trial**

In a pilot randomized controlled trial, we are testing the delivery of the multimodal intervention versus usual care control in patients with COPD and/or HF. In addition to study and intervention feasibility and acceptability measures, we will explore the mediating components of the mind–body exercise approach articulated in our conceptual model and assess clinically meaningful changes in self-efficacy and overall physical activity over the long term of 1 year. At Beth Israel Deaconess Medical Center, 42 participants with COPD and/or HF have been randomized in a 2:1 ratio to participate in the 1-year Mindful Steps program or to receive usual care (including verbal and written instructions to exercise).

**Study Population and Recruitment**

The inclusion criteria include the following: age over >40 years; clinical diagnosis of COPD, defined as either a ratio of forced expiratory volume in one second (FEV<sub>1</sub>) to forced vital capacity <0.70 or chest computed tomography evidence of emphysema, and/or clinical diagnosis of HF syndrome (with left ventricular systolic dysfunction or preserved ejection fraction, and New York Heart Association Class 1-3); medical clearance from a provider to participate in an exercise program; an active email account with the ability to check email at least weekly; and access to a computer with an internet connection, USB port, and Windows. Meanwhile, the exclusion criteria include the following: self-reported COPD or HF exacerbation in the previous 2 weeks; inability to ambulate; clinical signs of unstable cardiovascular disease; hypoxemia during 6-minute walk test (oxygen saturation <85% with supplemental oxygen); inability to collect at least 7 of 14 days of baseline step counts; current participation in a cardiac or pulmonary rehabilitation program.

To identify potential participants, we searched hospital registries, screened physician schedules, solicited direct physician referrals through email, and used targeted patient advertising (paper flyers, digital screen ads, in-person presentations) in primary care or specialty clinics at Beth Israel Deaconess Medical Center and pulmonary rehabilitation clinics in the community. This was followed by opt-in mailing and telephone outreach conducted by our research assistants.

**Randomization and Intervention Delivery**

As in prior studies, we used a 14-day run-in period of pedometer (Fitbit) use requiring at least 7 days of valid data (>200 steps per day) as a criterion for participants to be eligible for randomization. As the sample size is relatively small, we will use blocked randomization with varying block sizes and a computer-generated sequence of random numbers with a concealed and unpredictable allocation scheme.

Participants in the Mindful Steps intervention will receive the pedometer, access to the Mindful Steps website, and a schedule of in-person mind–body exercise classes as described in detail previously in this paper. All participants will continue pharmacological treatment and receive care through their usual providers. Participants in both groups will also receive an education handout from Harvard Health Publishing and Harvard Medical School titled “Walking for Health,” which includes information on the health benefits of walking, how to get started on a walking program, and specific walking workouts [104].

**Study Assessments**

Table 2 outlines study feasibility and intervention acceptability and adherence measures that are being tracked to inform a future larger trial.

<table>
<thead>
<tr>
<th>Outcomes and measures</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study feasibility</strong></td>
<td></td>
</tr>
<tr>
<td>Recruitment</td>
<td>Track recruitment rate by site and strategy (e.g., clinic visits, mailings, advertisement sites, involvement of physician, etc.)</td>
</tr>
<tr>
<td>Retention</td>
<td>Track retention rate with acceptable retention defined as &lt;20% dropouts</td>
</tr>
<tr>
<td><strong>Intervention acceptability</strong></td>
<td></td>
</tr>
<tr>
<td>Qualitative interview</td>
<td>Semistructured interview at 6 months (in person) and 12 months (by phone) to assess patient experience, understand helpful components of the multimodal intervention, and specifically explore themes related to conceptual model</td>
</tr>
<tr>
<td>Physical Activity Enjoyment Scale</td>
<td>8-item shortened version used to assess enjoyment of physical activity</td>
</tr>
<tr>
<td>[105]</td>
<td></td>
</tr>
<tr>
<td><strong>Intervention adherence</strong></td>
<td></td>
</tr>
<tr>
<td>MBE&lt;sup&gt;a&lt;/sup&gt; adherence</td>
<td>Track MBE class attendance, online video tutorial use, and logs of home practice</td>
</tr>
<tr>
<td>Web platform usage</td>
<td>Track logins to the website, clicks on education links, downloads and views of videos, and pedometer use</td>
</tr>
</tbody>
</table>

<sup>a</sup>MBE: mind–body exercise.
In addition, all participants will undergo testing to collect exploratory outcome measures at baseline, month 3 and 6 in-person, and month 9 and month 12 by phone. Table 3 outlines exploratory patient-centered outcomes including cognitive-behavioral and psychosocial measures, overall physical activity, disease-specific HRQL and symptoms, and physical function and exercise capacity.

At each time point, we will administer questionnaires and collect physical activity data. At in-person visits only, we will additionally perform the 6-minute walk test. All tests will be conducted by study staff who will be blinded to treatment assignment. We will also track potential adverse events and COPD- and HF-related exacerbations and hospitalizations at each testing visit to assess safety.

Qualitative interviews at 6 months and 12 months will further assess facilitators and barriers to participation and acceptability of the intervention and its multiple components. Additional questions will be guided by our conceptual model and explore themes of motivation, goal attainment, social support, self-reflection, and personal transformation.

Table 3. Patient-centered outcomes.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive–behavioral and psychosocial</strong></td>
<td></td>
</tr>
<tr>
<td>Self-Efficacy for Exercise Scale [106]</td>
<td>9-item validated scale from McAuley’s original barriers scale</td>
</tr>
<tr>
<td>Self-Efficacy for Managing Chronic Disease Scale [107]</td>
<td>6-item subscale from the Chronic Disease Self-Management Study</td>
</tr>
<tr>
<td>Intrinsic Motivation Inventory [108]</td>
<td>22-item validated scale with 4 subscales: interest/enjoyment, perceived competence, perceived choice, and pressure/tension</td>
</tr>
<tr>
<td>Patient Activation Measure [109]</td>
<td>13-item short form validated scale from Insignia Health</td>
</tr>
<tr>
<td>Medical Outcomes Study Social Support Survey [110]</td>
<td>20-item validated scale with 5 subdomains of tangible support, emotional support, informational support, affective support, and positive social interaction</td>
</tr>
<tr>
<td>Multidimensional Assessment of Interoceptive Awareness [111]</td>
<td>32-item validated scale with 5 subscales of body sensations, emotional reaction/attentional response, attention regulation, mind–body integration, trusting body sensations, assessing 8 domains of body awareness</td>
</tr>
<tr>
<td>CES-D(a) [112]</td>
<td>20-item validated scale; participants with CES-D score &gt; 20 will be referred back to primary provider for evaluation</td>
</tr>
<tr>
<td><strong>Physical activity</strong></td>
<td></td>
</tr>
<tr>
<td>Community Healthy Activities Model Program for Seniors Physical Activity Questionnaire [113]</td>
<td>41-item validated scale with leisure, household, occupational physical activity domains; allows estimation of weekly caloric expenditure</td>
</tr>
<tr>
<td>Pedometer [114]</td>
<td>Step counts measured by Fitbit pedometer; device worn during waking hours except when showering/bathing</td>
</tr>
<tr>
<td><strong>Disease-specific HRQL(b)</strong></td>
<td></td>
</tr>
<tr>
<td>St. George’s Respiratory Questionnaire [115,116]</td>
<td>Respiratory-specific extensively validated measure of HRQL; total score and subscales of activity, symptoms, and impact</td>
</tr>
<tr>
<td>Minnesota Living with HF(c) [117]</td>
<td>HF disease–specific extensively validated measure of HRQL</td>
</tr>
<tr>
<td>mMRC(d) Dyspnea scale [118]</td>
<td>5-item scale assessing breathlessness (part of BODE(e) index)</td>
</tr>
<tr>
<td><strong>Physical function and exercise capacity</strong></td>
<td></td>
</tr>
<tr>
<td>PROMIS(f) Physical Function, Fatigue [119]</td>
<td>Short form 7-10 items; from NIH(g) toolbox PROMIS</td>
</tr>
<tr>
<td>6-minute walk test [120]</td>
<td>Distance walked in 6 minutes as a measure of exercise capacity, performed according to ATS(h) guidelines; subjects will use supplemental oxygen if already prescribed oxygen during activity</td>
</tr>
</tbody>
</table>

\(a\)CES-D: Center for Epidemiological Studies Depression scale.

\(b\)HRQL: health-related quality of life.

\(c\)HF: heart failure.

\(d\)mMRC: Modified Medical Research Council.

\(e\)BODE: body-mass index, airflow obstruction, dyspnea, and exercise.

\(f\)Patient-Reported Outcomes Measurement Information System.

\(g\)NIH: National Institutes of Health.

\(h\)ATS: American Thoracic Society.
**Statistical Analysis**

For study feasibility and intervention adherence, we will use descriptive statistics to evaluate recruitment and retention rates, attendance at mind–body exercise classes, and adherence with the online intervention (web-platform usage). We will consider study retention successful if the retention rate is at least 80% and if attendance at in-person classes is at least 70% (17/25 classes). Mean scores on the Physical Activity Enjoyment Scale at each time point will be reported [105].

To further evaluate intervention acceptability, 6-month and 12-month qualitative interviews will be recorded and transcribed verbatim. Thematic analysis will be informed by grounded theory methods based on our semistructured questions. Transcripts will be independently coded by at least 2 separate investigators to identify emergent themes in an iterative process until thematic saturation is reached. We will particularly search for themes that may give insights into factors mediating a shift in self-efficacy towards longer-term maintenance of walking and overall physical activity.

For each patient-centered outcome, repeated measures analyses will be performed using SAS PROC MIXED to evaluate trajectories over the course of the 4 time points for each group and the differences between the groups in these trajectories at each time point. Planned paired comparisons and Cohen’s d effect sizes will also be computed to evaluate the magnitude of improvement between time points and group differences in change. These data will provide valuable information including what outcome domains and measures may demonstrate clinically meaningful changes in future models, the appropriate number of time points to use in a future clinical trial, and measures of variability within and between groups to inform the design of future studies. Measures considered sensitive to change will be those that indicate trends in improvement over time.

**Results**

Mindful Steps was funded in February 2017, approved by the institutional review board at Beth Israel Deaconess Medical Center in January 2017 and at the University of Michigan in May 2017. Phase I intervention development was completed in March 2019, phase II data collection began in April 2019, and completion is expected by August 2021. A total of 41 subjects have been enrolled.

**Discussion**

Through the integration of a web-based physical activity intervention with a mind–body video curriculum and in-person mind–body exercise, we created a novel, multimodal program with the goal of increasing long-term physical activity adherence for patients with COPD and HF. Guided by a conceptual model, the intervention aims to target multiple factors underlying behavior change that collectively act to enhance and internalize self-efficacy, which may lead to a longer-term, more sustainable shift in physical activity behavior. This pilot will provide valuable information needed to design a future clinical trial further assessing efficacy of this multimodal approach.

**Acknowledgments**

Grant support: This study was supported by an award from the National Center for Complementary and Integrative Health (NCCIH) at the National Institutes of Health (NIH) (R34AT009354). Dr. Yeh was supported by K24AT009465. Dr Wayne was supported by K24AT009282. Dr Rist was supported by K01HL128791. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NCCIH or the NIH.

**Conflicts of Interest**

PW is the founder and sole owner of the Tree of Life Tai Chi Center. PW’s interests were reviewed and are managed by the Brigham and Women’s Hospital and Partners HealthCare in accordance with their conflict of interest policy. No other authors have conflicts to declare. The Tree of Life Tai Chi Center did not participate in developing or administering the intervention for this study.

Multimedia Appendix 1

Peer-review report. [PDF File (Adobe PDF File), 121 KB - resprot_v10i4e27826_app1.pdf ]

**References**


Abbreviations

- COPD: chronic obstructive pulmonary disease
- HRQL: health-related quality of life
- HF: heart failure

Please cite as:

Litrownik D, Gilliam EA, Wayne PM, Richardson CR, Kadri R, Rist PM, Moy ML, Yeh GY
Development of a Novel Intervention (Mindful Steps) to Promote Long-Term Walking Behavior in Chronic Cardiopulmonary Disease: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2021;10(4):e27826
URL: https://www.researchprotocols.org/2021/4/e27826
doi:10.2196/27826
PMID:33913819

©Daniel Litrownik, Elizabeth A Gilliam, Peter M Wayne, Caroline R Richardson, Reema Kadri, Pamela M Rist, Marilyn L Moy, Gloria Y Yeh. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.
Uncovering the Bone-Muscle Interaction and Its Implications for the Health and Function of Older Adults (the Wellderly Project): Protocol for a Randomized Controlled Crossover Trial

Cassandra Smith1,2, MSc; Xuzhu Lin1, PhD; David Scott2,3,4, PhD; Tara C Brennan-Speranza5, PhD; Ahmed Al Saedi2,6, PhD; Alba Moreno-Asso1,2, PhD; Mary Woessner1, PhD; Ebrahim Bani Hassan2,6, PhD; Nir Eynon1, PhD; Gustavo Duque2,6, MD; Itamar Levinger2,7, PhD

1Institute for Health and Sport, Victoria University, Melbourne, VIC, Australia
2Australian Institute for Musculoskeletal Science, University of Melbourne and Western Health, Melbourne, VIC, Australia
3Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, VIC, Australia
4Department of Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia
5School of Medical Sciences and School of Public Health, Faculty of Medicine and Health, University of Sydney, New South Wales, Australia
6Department of Medicine-Western Health, The University of Melbourne, Melbourne, VIC, Australia
7Institute for Health and Sport, Victoria University, Melbourne, VIC, Australia

Corresponding Author:
Itamar Levinger, PhD
Institute for Health and Sport
Victoria University
PO Box 14428
Melbourne, VIC
Australia
Phone: 61 399195343
Email: itamar.levinger@vu.edu.au

Abstract

Background: Bone and muscle are closely linked anatomically, biochemically, and metabolically. Acute exercise affects both bone and muscle, implying a crosstalk between the two systems. However, how these two systems communicate is still largely unknown. We will explore the role of undercarboxylated osteocalcin (ucOC) in this crosstalk. ucOC is involved in glucose metabolism and has a potential role in muscle maintenance and metabolism.

Objective: The proposed trial will determine if circulating ucOC levels in older adults at baseline and following acute exercise are associated with parameters of muscle function and if the ucOC response to exercise varies between older adults with low muscle quality and those with normal or high muscle quality.

Methods: A total of 54 men and women aged 60 years or older with no history of diabetes and warfarin and vitamin K use will be recruited. Screening tests will be performed, including those for functional, anthropometric, and clinical presentation. On the basis of muscle quality, a combined equation of lean mass (leg appendicular skeletal muscle mass in kg) and strength (leg press; one-repetition maximum), participants will be stratified into a high or low muscle function group and randomized into the controlled crossover acute intervention. Three visits will be performed approximately 7 days apart, and acute aerobic exercise, acute resistance exercise, and a control session (rest) will be completed in any order. Our primary outcome for this study is the effect of acute exercise on ucOC in older adults with low muscle function and those with high muscle function.

Results: The trial is active and ongoing. Recruitment began in February 2018, and 38 participants have completed the study as of May 26, 2019.

Conclusions: This study will provide novel insights into bone and muscle crosstalk in older adults, potentially identifying new clinical biomarkers and mechanistic targets for drug treatments for sarcopenia and other related musculoskeletal conditions.


International Registered Report Identifier (IRRID): DERR1-10.2196/18777
Introduction

Background

Adults reach their peak muscle and bone mass in the third decade of life, after which an age-related loss of skeletal muscle and bone mass occurs [1,2]. Under certain conditions, for reasons that are not fully understood, this loss of bone mass (osteoporosis) and muscle (sarcopenia) is accelerated and, in some cases, occurs concurrently [3-5]. Emerging evidence suggests that this parallel and exponential loss of bone and muscle mass and strength is driven, at least in part, by bone and muscle crosstalk. The skeleton and skeletal muscle are closely linked anatomically, biochemically, and metabolically and modulate each other in endocrine and paracrine manners [6]. Many factors may be involved in this crosstalk, including genetics, changes in vitamin D and parathyroid hormone (PTH) levels, aging, increased levels of systemic and local inflammatory markers (ie, interleukin-6 [IL-6] and tumor necrosis factor), obesity and adipokines, mechanical loading, and altered hormones (ie, osteocalcin, resistin, and myostatin) [7,8]. The exact mechanisms involved in this crosstalk remain partially explored, although it has been proposed that undercarboxylated osteocalcin (ucOC) and possibly circulating osteoprogenitor (COP) cells may be mediators [6,9-12]. Serum total osteocalcin (tOC) is an osteoblast-specific secreted protein within the circulation that can be present in two major forms, as follows: γ-carboxylated osteocalcin (cOC) and ucOC lacking γ-carboxylation at one or more sites [13]. cOC, which is predominantly located in bone, is at least partly involved in bone mineralization, whereas ucOC has been shown to be involved in glucose metabolism—at least in mice—with new evidence suggesting a role in influencing muscle mass and strength [10,14-24]. Osteocalcin-deficient mice have reduced muscle mass and strength [17], and lower ucOC levels following hindlimb immobilization in rats are associated with reduced muscle mass and muscle force [25]. Treatment with ucOC can increase the cross-sectional area of the extensor digitorum longus, improve grip strength in mice, and stimulate myotube formation in C2C12 myoblast cultures in vitro [16]. In humans, we and others have shown that exercise increases serum ucOC levels and improves muscle metabolism and whole-body glucose control [10], most likely via increased insulin signaling protein levels within skeletal muscle, and that a decreased ratio of ucOC/tOC correlates with lower muscle strength in older women [19]. However, the effect of acute exercise on ucOC in older adults remains unknown; in particular, the role of ucOC in human myotubes and its association with muscle function parameters (ie, strength and mass) remain unclear.

Exercise causes a series of physiological responses in the bone and skeletal muscle, improving glucose regulation and insulin sensitivity and, importantly, promoting pro-osteogenic factors, including increasing bone formation biomarkers such as osteocalcin [26-31]. Exercise is a known nonpharmacological approach to improving bone health, reducing the risk of osteoporosis, and, importantly, concomitantly improving muscle function [32-35]. Thus, exercise represents an efficacious approach in older persons to reduce age-associated alterations related to sarcopenia, which currently has no available drug treatments. Evidence suggests that various mechanical factors including exercise stimulate the differentiation of mesenchymal stem cells (MSCs) into osteoblasts [36,37]. COP cells circulate within the blood and are MSC-like with osteogenic potential and a precursor for the osteoblastic lineage and potentially osteocalcin [11]. It remains unknown whether exercise, with stimuli promoting pro-osteogenic factors (ie, high load bearing resistance exercise [RE], impact, and jumping exercise), can mitigate the aging process in skeletal muscle and bone through its effect on COP cell levels and therefore osteocalcin [38]. Even a single acute bout of exercise (ie, aerobic exercises [AERs] and REs) elicits positive effects on the bone endocrine and biomarker response and can increase ucOC and insulin sensitivity [9,10,39-42]. However, the exact mechanism by which ucOC influences skeletal muscle function, strength, and metabolism in older adults remains unclear.

Objectives

The primary objective of this study is to determine the change in circulating ucOC following acute exercise interventions between older adults with low and high muscle function (stratified based on leg muscle quality [LMQ]; see the Lower Limb Maximal Strength and LMQ section for the LMQ equation). To extend the current knowledge of ucOC in humans, and as an adjunct to this study, we will also perform in vitro experiments on cultured primary myotubes to uncover the mechanistic pathways of action of ucOC. In addition, as a secondary objective, we will aim to quantify the lineage of COP cells before and after exercise, as COP cells can potentially act as a regeneration and antiaging inducible factor and a precursor for osteocalcin.

Hypotheses

We hypothesize that older adults with lower muscle function, compared to those with normal or higher muscle function, will (1) be characterized by lower levels of circulating ucOC and those with lower ucOC will be associated with poorer glucose control and (2) be characterized by abnormal skeletal muscle signaling (muscle hypertrophic or atrophic pathways). Both AER and RE will increase ucOC; however, we hypothesize that this will be to a greater degree in those with lower muscle function. We will test these hypotheses at baseline and after acute exercise. This project has the potential to identify novel biomarkers for interactions between bone and muscle, with implications for future drug targets or clinical interventions and the management of those with or at risk of sarcopenia or reduced muscle function.
**Methods**

**Design**

This study is a randomized controlled crossover trial (Figure 1) approved by the Melbourne Health (MH) Human Research Ethics Committee (reference number: 2017/08) and is registered with the Australian New Zealand Clinical Trials Registry (trial number: ACTRN12618001756213). The trial is a multicenter clinical trial conducted at the Institute for Health and Sport, Victoria University, Melbourne, Victoria, Australia, and the Australian Institute for Musculoskeletal Science (AIMSS) in Western Health, St Albans, Victoria, Australia. The trial will be conducted in accordance with the Helsinki Declaration, and reporting of the study will adhere to the CONSORT (Consolidates Standards of Reporting Trials) guidelines [43,44].

**Figure 1.** Study design. 1RM: one-repetition maximum; AER: aerobic exercise; pQCT: peripheral quantitative computed tomography; RE: resistance exercise.

**Participants**

Men and women aged 60 years or older will be recruited. Women will be required to be a minimum of 12 months postmenopause; this is because of the potential alteration in hormones that occur during perimenopause, which can interact with or affect the specific project outcomes of this study. The inclusion and exclusion criteria are listed in Textbox 1. Additional study exclusions will include the inability to provide informed consent independently for safety reasons (particularly as we take some invasive measures) and an inability to understand English, as this may potentially be a safety concern if unable to communicate during visits that include maximal exertion testing and acute exercise bouts.

**Textbox 1.** Study eligibility.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males and females aged 60 years</td>
</tr>
<tr>
<td>Females &gt;12 months postmenopause</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any fractures within the previous 3 months</td>
</tr>
<tr>
<td>Have begun a new osteoporotic treatment within the previous &lt;3 months or have begun taking antiresorptive medications within the previous &lt;3 months</td>
</tr>
<tr>
<td>Have diabetes mellitus or are taking hyperglycemic medications</td>
</tr>
<tr>
<td>Any hematological, myelodysplastic, or myeloproliferative disorder</td>
</tr>
<tr>
<td>Any bone malignancy</td>
</tr>
<tr>
<td>Taking warfarin of vitamin K supplementation or restriction</td>
</tr>
<tr>
<td>BMI≥40 kg/m²</td>
</tr>
<tr>
<td>Engagement in a resistance exercise regime for more than 2 sessions per week</td>
</tr>
</tbody>
</table>

**Recruitment**

Prospective participants will be recruited using advertisement flyers. These will be displayed at Western Health sites (Sunshine and Footscray Hospitals, Victoria, Australia) and provided for use within the general community and other media outlets. Interested participants will self-initiate contact with the research team via email or phone. Those interested will be screened...
against the inclusion and exclusion criteria. Eligible participants will be provided with information for the participants and participant-informed consent forms. A physical examination and an approval to participate in the study will be required from the patients’ physician. Please refer to Multimedia Appendix 1 for the schedule of enrollment, interventions, and assessments.

Initial Screening

Summary of Initial Screening

The initial screening will be used for clinical characterization of the volunteers as well as for bone and muscle quantification and quality. It includes 2 separate visits of 3-4 hours’ duration (visits 1 and 2; Figure 1), performed in any order and up to 14 days apart. Both visits will be performed in the morning and following an overnight fast. The measures obtained during these visits are explained in detail below.

Bone and Muscle Health

Dual Energy X-Ray Absorptiometry

Body composition and bone mineral density (BMD) will be assessed using a dual energy x-ray absorptiometry (DXA) scanner (Hologic, Horizon A, software version 5.6.0.4). Total BMD as well as the neck of the femur and lumbar spine BMD will be assessed. Lean body mass and fat mass will also be assessed. The DXA scan will ideally be performed in the morning following an overnight fast by experienced personnel. This will be completed by the AIMSS.

Bone Microarchitecture and Fat Infiltration

Peripheral quantitative computed tomography (pQCT; Stratec XCT3000, Stratec Medizintechnik GmbH) will be used to quantify muscle and bone mass, density, and adipose infiltration at the nondominant forearm and foreleg [45,46].

Single 2.5-mm transverse scans will be obtained at 4% and 66% of tibial length (measured form the palpable tip of medial malleolus) and 4% and 66% of the radial length (from the radial condyle), with a voxel size of 0.4 mm. All pQCT scans will be acquired and analyzed by an experienced operator, and the device will be calibrated on the scan date using the manufacturer’s phantom. Calf and forearm muscle cross-sectional areas (mm²) and densities (mg/cm³) will be determined using the manufacturer’s algorithms and software (version 6.2). The calf intramuscular adipose tissue cross-sectional area (cm²) will be quantified as previously described [47]. Trabecular and cortical bone densities and structure will be assessed at the relevant regions of interest. All imaging will be performed by an appropriate expert (radiographer).

Blood Sample

Quantification of Bone Remodeling and Cardiometabolic Biomarkers

Beta-isomerized C-terminal telopeptide (a bone resorption marker) and procollagen 1 N-terminal propeptide (a bone formation marker) will be quantified using a Roche Hitachi Cobas e602 immunoassay analyzer, according to the manufacturer’s guidelines. Hormones (PTH), lipids, glucose and insulin, inflammation markers (C-reactive protein and serum IL-6), and potentially other cardiovascular or health markers will be analyzed according to standard hospital procedures.

Genotyping and Target Genetic Variants Analyses

We will target either candidate gene variants [48,49] or Genome-Wide Association–based variants previously related to skeletal muscle and bone health [50,51]. Genomic DNA will be extracted from residual blood samples from Becton Dickinson (BD) Vacutainer EDTA tubes using the MagSep Blood gDNA kit (0030 451.00, Eppendorf) or GeneJET Genomic Whole Blood DNA Purification Kit (#K0781 Thermo Scientific). Gene variants will be determined using the TaqMan SNP assay (Applied Biosystems, Thermo Fisher Scientific) by QuantStudio 7 Flex (Applied Biosystems, Thermo Fisher Scientific). Genotyping will be replicated in another independent institute, as previously described [52,53], to validate the results.

Muscle Function and Strength

Grip Strength and Gait Velocity

Grip strength will be measured using a hand dynamometer; a result of <20 kg for women and <30 kg for men will identify low muscle strength [54,55]. A 4 m gait velocity assessment will be performed by using the instrumented walkway, which has an acceleration (GAITRite), and by timing with a stopwatch, and reduced physical function will be determined as <80 cm/second. Both the grip strength and gait velocity thresholds noted are accepted as a measurement of sarcopenia [54,55] and will form the definition in this study.

Lower Limb Maximal Strength and LMQ

Participants will perform a one-repetition maximum (1RM) test on a leg press. This will be performed twice, with the first visit serving as familiarization. 1RM is defined as the heaviest weight lifted once with the proper technique and without compensatory movements [56]. Results from this study will guide appropriate prescription for the acute RE session.

LMQ, an estimate of specific force, has been shown to decrease with age and is described as the amount of force a muscle group can produce per unit of muscle mass [57]. We will calculate LMQ as follows:

\[
LMQ = \frac{\text{leg strength (kg)}}{\text{(left leg lean mass [kg] + right leg lean mass [kg])}} \times 100
\]

Leg strength will be defined as the participants’ 1RM, and leg lean mass will be obtained from the DXA assessment.

Physical Performance Test

Participants will complete a physical performance test (PPT), adapted from Levinger et al [59], and will include 4 functional mobility tasks: (1) a gait velocity assessment (described earlier), (2) timed up and go test, (3) stair climbing power (SCP), and (4) stair descending. All tests will be scored in time (seconds).

The timed up and go test is a simple performance-based assessment that requires minimal equipment, including a standard arm chair (approximately 46 cm), a 3-m walkway with a floor mark, and a stopwatch (time, seconds). It is performed as time (seconds) taken to rise from a seated position, walk 3
m, turn, walk back to the chair, and then sit. The SCP consists of a rapid ascent of 10 stairs and is calculated as follows:

\[
P = \text{force} \times \text{velocity}^2\]

[60]

Velocity is calculated as the vertical distance of the stairs divided by the time it takes to ascend the stairs. Force is calculated as the participants’ body weight multiplied by acceleration due to gravity (9.8 m/s²). The stair descent will be the time to safely descend 10 stairs. The rest between ascent and descent will be 45 seconds. Participants will undergo 4 attempts on each task, and the best time will be recorded for each task. The PPT score will be the sum of the fastest times recorded for each test.

**Aerobic Capacity and Vascular Health**

Peak oxygen consumption will be assessed on a cycle ergometer with the initial intensity beginning at 10-30 W and increasing by 10-30 W×minute⁻¹ according to participant ability. Participants will be monitored by a 12-lead electrocardiogram (Mortara, X-Scribe II). Oxygen consumption for each 15-second interval will be measured by gas exchange analysis (BreezeEx, version 3.02, Medical Graphics Corporation), with routine calibration of gas concentrations and flow before each test. The test will be terminated according to participants’ self-reported fatigue perception reaching a predetermined level (using the Borg scale, ratings of perceived exertion [RPE]=17) or clinical signs or symptoms [61]. Blood pressure will be monitored at baseline, regular intervals (each stage), and post exercise using a manual sphygmomanometer, and heart rate will be monitored via the 12-lead electrocardiogram.

Vascular endothelial function will be assessed by brachial artery flow–mediated dilatation, used in clinical trials as a reproducible method to assess endothelial function [62,63]. Vascular stiffness will be assessed by noninvasive measures of pulse wave velocity (simultaneous comparison of carotid and femoral arterial pulses) and pulse wave analysis (pulsations recorded at the brachial artery to produce central aortic pressure waveforms) using applanation tonometry (SphygmoCor EXCEL system V1, AtCor Medical) [64].

**Questionnaires and Lifestyle Behaviors**

**Physical Activity Log**

Participants will complete a lifestyle behavior and physical activity log. This log has been developed for the purpose of this study and will have questions related to average sleep cycles and normal physical activity levels on weekdays versus weekends (stratified into moderate, hard, and very hard activities). The physical activity component includes consideration for activities of daily living and structured exercise, with examples provided.

**Dietary Behavior**

A 3-day dietary log will be given to participants on their first visit, to be returned on visit 2 for investigators to analyze normal dietary behaviors. Participants are encouraged to eat normally while they are recording (ie, not to adjust food quantities) and are instructed to complete the dietary log on 2 weekdays and 1 weekend day (consecutively). This log also requests a timed record of physical activity behaviors, including the time and intensity, the time at which food and drinks are ingested, and the time and quantity of medications and supplements.

**Falls Risk Questionnaire**

The Falls Risk for Older People in the Community (FROP-Com) was developed by the National Ageing Research Institute as a modified version of the Falls Risk for Hospitalized Older People for better utility in the community [65]. The FROP-Com is simple, takes only 10-15 minutes to complete, is low cost, requires no equipment, and can be administered by any health professional. It is a comprehensive fall risk assessment, covering 13 risk factors for falls set out in 26 questions with dichotomous or ordinal scoring (from 0 to 3). The overall score is indicative of fall risk, with the total score ranging from 0 to 60, with higher scores indicating greater risk. The tool has demonstrated good reliability and has a moderate capacity to predict falls [65].

**Mini Nutritional Assessment Questionnaire**

The Mini Nutritional Assessment (MNA) is a widely used tool for assessing nutritional status in older adults. It is simple to administer, low cost, and validated, with high sensitivity, specificity, and reliability. The MNA classifies the interviewee as well nourished (score≥24), at risk of malnutrition (score between 17 and 23.5), or malnourished (score<17). The MNA also correlates with clinical assessments and objective measures, such as albumin, BMI, triceps skinfold, caloric intake, and vitamin status, and low scores are related to the incidence of clinical events and mortality [66-69].

**Charlson Comorbidity Index Questionnaire**

The Charlson Comorbidity Index (CCI) is a validated measure of 1-year mortality risk and burden of disease and is used in clinical research to understand the influence of comorbidities and predict outcomes [70-72]. In clinical practice, the CCI assists with the stratification of patients into subgroups based on disease severity to assist with targeted models of care and resource allocation. The CCI includes 17 comorbidities (with 2 subgroups for diabetes and liver disease) that are weighted between 17 and 23.5), or malnourished (score<17). The MNA also correlates with clinical assessments and objective measures, such as albumin, BMI, triceps skinfold, caloric intake, and vitamin status, and low scores are related to the incidence of clinical events and mortality [66-69].

**Randomization**

Following baseline assessments, participants will be randomized into the acute intervention to explore the characteristics of older adults (by sex) with low or high muscle function (Figure 1). Participants will be randomized individually by a researcher external to this project (they will have no contact with the participants before or during the trial). This person will also have no intellectual or personal investment in the study design, data collection, or outcome. The order of the 3 conditions for each participant (AER, RE, or control) will be randomized using a sealed envelope method (block allocation) to prevent carryover effects between conditions. The envelopes will be stored separately in a locked cabinet, and each envelope will contain 3 pieces of paper that will state “AER,” “RE,” and “CON.” These pieces of paper will be folded to reduce their transparency.

https://www.researchprotocols.org/2021/4/e18777
**Study Intervention**

**Acute Intervention**

Participants will complete the acute intervention (visit 3, 4, and 5) up to 14 days after completing the screening assessments and will complete the AER, RE, and CON conditions (Textbox 2) in a randomized order (described below). These visits will include blood sampling and optional skeletal muscle biopsies. Participants can elect to have none, 1 (at rest, for a baseline measure), or 4 biopsies (1 at rest for baseline and 1 following each condition). Each testing visit is approximately 3 hours in duration, including the 30-minute intervention (exercise or rest), and visits will be performed approximately 7 days apart, accounting for washout.

Textbox 2. Acute intervention.

**Description of interventions**

- **Aerobic exercise:** Performed on the cycle ergometer for 30 minutes at an intensity corresponding to 70%-75% of peak heart rate; this is based on data obtained from the exercise capacity assessment. Intensity will be adjusted every 5 minutes to maintain the desired heart rate range.

- **Resistance exercise:** The protocol is as we have previously performed [10] and includes 30 minutes of strength and power exercises at intensities corresponding to 70%-75% of the predetermined one-maximal repetition based on the individual’s test results. Leg press will be performed as 5 sets of 10 rapidly concentric (as fast as possible) and slow eccentric (4 seconds) repetitions. Recovery between sets and exercise will be 2 minutes. Participants will also perform jumping sequences as 5 sets of 10 jumps (jumping as high as they can 10 times without stopping). Power training is effective to increase muscle strength and bone density and is safe for older adults [73-75].

- **Control:** This session will include 30 minutes of supine bed rest.

All testing visits will be monitored and supervised by accredited exercise physiologists (AEPs), who will follow the structured protocol as dictated for that particular session (AER, RE, and CON). The AEP will also monitor signs and symptoms in response to exercise training and will record Borg RPE, blood pressure, and heart rate at frequent time points. Any adverse signs and symptoms will be documented, including feelings of fatigue, soreness, light-headedness, and any injuries. Blood sampling and intravenous cannulation will be performed by personnel who are experienced in the technique, and muscle biopsy will be performed by an experienced medical physician.

**Control Procedures**

For testing visits 3, 4, and 5, participants will arrive at the laboratory between 7 AM and 8 AM following an overnight fast and with abstinence from exercise or reduced general activity (ie, heavy to moderate activities of daily living) in the preceding 24 hours and for all follow-up sessions. These sampling procedures will be followed at all visits to account for circadian or diurnal rhythms [76]. Participants may be requested to abstain from particular medications (eg, aspirin), as advised by the medical doctor, if electing for a muscle biopsy.

To assist with adherence to study protocols, participants will be monitored via regular communication with the study coordinator on the days preceding each study visit. As a general consideration for participation in this study, participants will be encouraged not to alter their current physical activity levels, exercise habits, or dietary intakes for the entirety of the study. Participants are asked to report whether there have been any alterations in medications throughout the study, as we request that all medication interventions are stable for at least more than three months.

**Biospecimen Sampling Protocols**

On arrival and following supine rest (approximately 15 minutes), a cannula will be inserted into the antecubital vein, and a baseline (resting) blood sample (40 mL; biopsy, if consented) will be obtained. These baseline (resting) samples will be obtained on the first visit only, serving as the baseline for all other visits thereafter. Four additional blood samples following the 30-minute acute intervention will be collected immediately after the intervention (0-minute time point) and at 30, 60, and 120 minutes postintervention (total of 100 mL) to observe changes in tOC, ucOC, COP, and other measures. If elected, a postintervention biopsy will be conducted at the 60 minutes time point. At all time points, blood samples will be collected into EDTA and serum-separating tubes vacutainers for the appropriate collection of serum or plasma. Following 10-minute clotting time, samples will be centrifuged for 10 minutes at 4°C and immediately transferred to long-term storage at ~80°C in 2 mL aliquots for later analysis.

**Outcome Measures**

**Primary Outcomes**

The primary outcome for this study is the peak change in circulating levels of ucOC from baseline compared with postacute exercise blood sampling time points (0, 30, 60, and 120 minutes) following the 3 acute interventions (AER, RE, and CON) between the low muscle function and high muscle function groups.

**Secondary Outcomes**

The secondary outcomes for this study are (1) the difference in protein content related to atrophic and hypertrophic protein signaling at baseline between the low muscle function and high muscle function groups and (2) the difference in protein signaling (protein content) from baseline and compared with the postmuscle sampling timepoint (60 minutes) following the 3 acute interventions (AER, RE, and CON) between the low muscle function and high muscle function groups.

**Data Collection and Analysis**

**Quantification of Osteocalcin**

Serum tOC will be analyzed as described previously [9,10,19,39]. In brief, tOC will be measured using an automated immunoassay (Elecsys 170; Roche Diagnostics). Serum ucOC will be measured by the same immunoassay after absorption of
cOC on 5 mg/mL hydroxyapatite-apatite slurry, following the method described by Gundberg et al [77].

**Quantification of COP Cells**

COP cell analysis will be performed as described previously [78,79]. In brief, peripheral blood samples (20 mL) will be collected (EDTA tubes) and processed for Ficoll-based gradient separation, and approximately $5 \times 10^6$ peripheral blood mononuclear cells (PBMCs) will be obtained. Approximately $1 \times 10^6$ PBMCs will be resuspended in fluorescence activated cell sorting (FACS) buffer, followed by a 10-minute blocking with fragment crystallizable receptor blocking reagent (BD Biosciences). Staining will then be performed with a viability marker (30 minutes), followed by washing (×2) with phosphate buffered saline (containing 5% fetal calf serum). Cells will be incubated with mouse anti-human CD45-Pacific Blue, CD3-PerCP, and CD19-APC (40 minutes). Staining of intracellular components will be permeabilized with Cytofix/Cytoperm (BD Biosciences) according to the manufacturer’s instructions, followed by incubation with mouse anti-human osteocalcin-phycocerythrin at 4 °C (40 minutes), and then washed with Perm or wash buffer (×2).

**Flow Cytometry**

Cells will be analyzed using a BD FACS Canto II. FACS DiVa software will be used to analyze 50,000 total events for each sample and for the fluorescence minus one (FMO) controls. A total of 3 lasers and 8 different photomultiplier tube (PMT) channels will be used for the 6-color staining panel. Two FMOs will contain fluorochromes, except for the one to be controlled for. Compensation beads will be used to set the compensation controls for each fluorochrome. The PMT voltage values for fluorochromes will be set for each cell type based on the compensation controls. Doublet discrimination will be applied, and viability will be assessed by negative staining using the Live/Dead stain. Offline analysis will be performed using FlowJo analytical software (Treestar).

**Gating Strategy**

Cells will be gated for size, shape, and granularity using forward and side scatter parameters, as previously described by our group [78,79]. Briefly, serial gating steps will be applied to quantify cellular populations. First, dead cells will be excluded, and then, a region will be set to encompass lymphocyte-, monocyte-, and granulocyte-enriched areas, followed by doublet discrimination, T cell (CD3) and B cell (CD19) elimination. For COP cells, after gating on live single mononuclear cells (on forward and side scatter plots), the CD45 and osteocalcin double positive cells will be calculated. Cutoff points to assign antigen positivity will be performed against matching FMO controls. The use of FMO significantly increases the sensitivity and specificity of analysis, as they effectively minimize the effect of nonspecific antibody binding and cell-specific autofluorescence. The gating quantification will be performed twice for the accurate quantification of the percentage of COP cells.

**Muscle Sampling Protocol and Analyses**

**Summary of Sampling Procedure**

If elected, the muscle samples (1 or 4) will be taken from the vastus lateralis (approximately 150 mg) under local anesthesia (xylocaine 1%) by percutaneous needle biopsy technique, modified to include suction [80]. Excised tissue will be snap frozen in liquid nitrogen and stored at −80°C for later analysis. Proteins involved in muscle degradation and hypertrophy (ie, anabolic and catabolic pathways; ubiquitin-proteasomal, autophagy-lysosomal, and caspase-3–mediated proteolytic pathways) as well as glucose uptake will be assessed, as described previously [81-84].

**Protein Extraction and Western Blotting**

All muscle samples (baseline and postintervention samples) will be used to analyze the content and activation of signaling proteins involved in muscle degradation and hypertrophy by using western blotting, as described previously [84-86]. Western blotting is a method commonly used to detect and analyze the abundance and posttranslational modifications (such as phosphorylation) of proteins. Briefly, muscle samples will be homogenized in a radioimmunoprecipitation assay buffer using a TissueLyzer (QIAGEN). Then, proteins in the lysate will be separated based on protein molecular weight via gel electrophoresis. Proteins will be subsequently transferred to a polyvinylidene fluoride membrane where specific proteins can be probed using specific antibodies. Finally, signals generated through electrophoresis will be detected and analyzed using ChemiDoc Imaging Systems (Bio-Rad Laboratories).

**Primary Skeletal Muscle Cell Culture**

A portion of the muscle obtained at rest (baseline) will be used for cell culture for future molecular analyses [87]. This will be established according to the method described by Blau and Webster [88] and by Gaster et al [89] and previously detailed by McAinch et al [90]. Briefly, muscle samples (50-100 mg) will be washed, minced, and enzymatically dissociated with trypsin. Cells will be cultured in a coated flask with extracellular matrix, and the growth medium (α-minimal essential medium (α-MEM)+10% fetal bovine serum+0.5% penicillin-streptomycin+0.5% antifungal) will be changed every other day until they reach 80% confluence. Then, satellite cells will be selected using CD56+ magnetic microbeads (Miltenyi Biotec) and transferred to bigger flasks coated with extracellular matrix to increase cell number (up to 4 passages). Once the cells reach 80% confluence, they will be differentiated using a differentiation medium (α-MEM+2% horse serum+0.5% penicillin-streptomycin+0.5% antifungal) for 5-6 days. Before the experimental treatment, cells will be starved for 1 hour in a serum-free medium (α-MEM+0.5% penicillin-streptomycin+0.5% antifungal).

In vitro treatment with ucOC at concentrations of 0 ng/mL, 30 ng/mL, and 100 ng/mL in serum-free medium for 60 minutes and 24 hours in the presence or absence of insulin (100 nM for the last 15 minutes) for the determination of glucose uptake (2-deoxy-D-[3H] glucose) and western blotting will be performed. Analysis of targeted proteins (described above) will
be performed as described previously [39,91]. The dose response is important because the physiological effect of ucOC in muscles from mice and humans may be different, and the concentrations used are all physiologically relevant.

**Participant Retention and Withdrawal**

Once a participant is enrolled into the trial, the study coordinator will keep in contact with him or her for the entire study period. We estimate that the dropout rate in this population will be approximately 10%. Participants may withdraw from the study at any given time. The investigators or medical staff may also withdraw participants from the study due to safety or medical concerns.

**Statistical Analysis and Determination of Sample Size**

The primary endpoint for this study is the change in ucOC levels from baseline to the peak, postintervention sampling time point. The analysis will include a comparison of changes in ucOC levels in response to each intervention from baseline to postintervention sampling time points (0, 30, 60, and 120 minutes) between the low muscle function and normal or high muscle function groups, by using repeated measures analysis of variance (ANOVA). Comparisons of multiple means will be examined using a 2-factor (exercise type x time point) repeated measures ANOVA. For all significant interaction and main effects, a priori comparisons of means (baseline vs all postexercise time points) will be conducted using the Fisher least significant difference test ($P < 0.05$). Multivariable regression models will be used to determine associations between selected measurements, adjusting for BMI and sex. Data will be analyzed using the Statistical Package for the Social Sciences, version 22 (SPSS Inc), and statistical significance will be declared at $P < 0.05$.

In 10 postmenopausal women, we previously reported that the change in ucOC levels following exercise is approximately 9% [39]. We will recruit 54 participants (equal number of men and women) who will be dichotomized as low muscle function versus high muscle function (27 per group). After adjusting for a loss to follow-up rate of 10% (5.4/54), this sample size will be large enough to detect an estimated 4% difference in changes in ucOC levels (SD 6%) between groups with a type I error rate of 5%, type II error rate of 20%, and power of >80% (G*Power 3.1.9.2 for Windows) [92].

**Data Monitoring**

**Data Management and Monitoring**

Details of the procedures for data management have been reviewed and approved by the MH Human Research Ethics Committee and can be located via study reference, 2017/08. A trial management group (TMG) has been formed to manage potential risks and for structured oversight of the trial [93]. The type of oversight that this TMG will provide includes regular meetings to review individual safety reports and data relating to quality, protocol adherence, and participant retention rates. The TMG committee will include the principal investigator; individuals responsible for the daily running of the trial, including the trial coordinator; and an appointed independent member.

All electronic data will be stored on password-protected computers. Hard copies of any data will be kept in a locked filing cabinet in a secure office. All data collection tools and questionnaire data will be deidentified.

**Harms**

All adverse events associated with the study, or occurring during study participation, will be recorded. All adverse events will be reported to the TMG and ethics committees (MH and Victoria University) with strategies to reduce the risk for future events. The ethics committees have the power to pause or even stop the research in the case of a severe adverse event. The study personnel will monitor the clinical signs and symptoms of dyspnea, shortness of breath, nausea, faint-headedness and light-headedness, signs of inflammation, or infection throughout the study period.

**Auditing**

The TMG will meet and run an internal audit of the trial at regular intervals annually. The principal investigator, JL, is responsible for the overall conduct and preservation of the integrity of this trial and has extensive experience as a lead research investigator in numerous human clinical trials.

**Ethics and Dissemination**

This study and its protocols and data collection tools have been reviewed and approved by the MH Human Research Ethics Committee (ref approval number: HREC/17/MH/335; local project number: 2017/208), and local ethical approval has been confirmed at Victoria University as a mirror approval of MH. Any modifications to the study objectives, procedures, protocols, data collection tools, and study personnel will require a formal amendment from the ethics committee. All protocols related to consenting procedures, data collection and access to participant data, procedures for maintenance of participant confidentiality, and plans for dissemination of study results can be found in the reviewed and approved documents of trial reference 2017/208.

**Results**

The trial is active, with participant recruitment and intervention delivery currently ongoing (MH Human Research Ethics Committee ref approval number: HREC/17/MH/335; Western Health Sunshine Hospital local project number: 2017/208; protocol version number: 6 25/06/2018). Recruitment for this trial began in February 2018, and 38 participants have completed the study as of May 26, 2019. The results of this study will be published throughout the trial, and the main study findings are expected to be published by June 2021.

**Discussion**

Previous research suggests that crosstalk exists between the skeleton and skeletal muscle; however, this crosstalk has not been fully described or clearly elucidated, particularly in humans. In addition, a lack of physical activity accelerates the widespread cellular and molecular changes induced by aging, resulting in an increased prevalence of many chronic diseases [94]. Detecting the age-related conditions associated with inactivity and early intervention are essential for reducing the...
economic burden of aging on the health care systems worldwide. The development of affordable and universally accessible ways to prevent chronic disorders, such as tailored exercise programs, in combination with the development of robust blood biomarkers, will considerably improve the ability to predict and detect chronic diseases and reduce the health and economic burden caused by aging in a cost-effective manner [95].

This project is designed to uncover a novel crosstalk pathway between bone and muscle in older adults via ucOC. Current evidence from predominately cross-sectional studies suggests that osteocalcin, via ucOC, in humans may be associated with muscle function [10,19]. Evidence from animal and preclinical studies is encouraging, indicating a promising role for ucOC in improving muscle metabolism and function [16,17,25]. However, the roles of ucOC in humans and its relationship with muscle function and metabolism remain unknown. Evidence suggests that COP cells have a dynamic capacity to mobilize to sites of fracture repair and have the capacity to be upregulated under varying pathological or physiological bone forming processes, such as puberty and fracture [79,96-102]. However, it is unknown whether exercise stimulus, with osteogenic capacity, can increase the COP cell population and upregulate tOC and therefore ucOC.

The proposed project aims to overcome this gap by characterizing ucOC levels in older adults with a spectrum of muscle functions and in response to an acute exercise intervention. Importantly, we will investigate this bone and muscle crosstalk by determining the associations between the parameters of muscle function (ie, muscle signaling, muscle mass, and muscle strength) and ucOC. In the future, we plan to directly assess this association at a cellular level in human primary myotubes (those prepared in this study) to determine the direct effects of ucOC on muscle protein signaling and glucose uptake. The results of this study will provide a greater understanding of skeletal muscle metabolism and the crosstalk between muscle and bone in the older adult population. We aim to establish ucOC as a biomarker for muscle function and bone and muscle crosstalk in older adults to target potential mechanisms for future therapeutic studies. We also aim to advance the development of personalized clinical exercise guidelines for sarcopenia and other musculoskeletal conditions.

Acknowledgments
This work is supported by research funds of the primary investigators, IL and GD, and the PhD (Doctor of Philosophy) funds of CS. This trial has been awarded the Tom Penrose Community Service Award from Exercise and Sport Science Australia, a seed grant from AIMSS, and a Diabetes Australia Research Program Grant from Diabetes Australia. The grants awarded to this trial did not have any role in or contribution to the design of the study and the collection, analysis, or interpretation of the data; the grants awarded to this trial will not contribute to the writing of any associated manuscripts.

Authors’ Contributions
CS participated in the trial design, coordination and intervention delivery of the trial, and collection of data at Victoria University and Western Health and drafted the manuscript. IL participated in the design of the study, contributed to data collection at Victoria University and Western Health, and helped draft the manuscript. GD participated in the study design, was involved in the trial as an advising medical physician, and reviewed the manuscript. The remaining authors, XL, DS, TB, AA, AM, MW, EB, and NE, contributed or specialized in specific expert techniques; and CS, IL, GD, XL, DS, TB, AA, AM, MW, EB, and NE read and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Schedule of enrollment, interventions, and assessments.

References


Abbreviations

1RM: one-repetition maximum
AEP: accredited exercise physiologist
AER: aerobic exercise
AIMSS: Australian Institute for Musculoskeletal Science
ANOVA: analysis of variance
BD: Becton Dickinson
BMD: bone mineral density
CCI: Charlson Comorbidity Index
cOC: γ-carboxylated osteocalcin
COP: circulating osteoprogenitor
dXA: dual energy x-ray absorptiometry
FACS: fluorescence activated cell sorting
FMO: fluorescence minus one
FROP-Com: Falls Risk for Older People in the Community
IL-6: interleukin-6
LMQ: leg muscle quality
MH: Melbourne Health
MNA: Mini Nutritional Assessment
MSC: mesenchymal stem cell
PBMC: peripheral blood mononuclear cell
PMT: photomultiplier tube
PPT: physical performance test
pQCT: peripheral quantitative computed tomography
PTH: parathyroid hormone
RE: resistance exercise
RPE: ratings of perceived exertion
SCP: stair climbing power
TMG: trial management group
tOC: total osteocalcin
ucOC: undercarboxylated osteocalcin
α-MEM: α-minimal essential medium
Uncovering the Bone-Muscle Interaction and Its Implications for the Health and Function of Older Adults (the Wellderly Project): Protocol for a Randomized Controlled Crossover Trial

JMIR Res Protoc 2021;10(4):e18777
URL: https://www.researchprotocols.org/2021/4/e18777
doi:10.2196/18777
PMID:33835038

©Cassandra Smith, Xuzhu Lin, David Scott, Tara C Brennan-Speranza, Ahmed Al Saedi, Alba Moreno-Asso, Mary Woessner, Ebrahim Bani Hassan, Nir Eynon, Gustavo Duque, Itamar Levinger. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 09.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

Electronic Health Record–Embedded Individualized Pain Plans for Emergency Department Treatment of Vaso-occlusive Episodes in Adults With Sickle Cell Disease: Protocol for a Preimplementation and Postimplementation Study

Lingzi Luo, MPH, MSW; Allison A King, MD, PhD, MPH; Yvonne Carroll, JD, RN; Ana A Baumann, PhD; Donald Brambilla, PhD; Christopher R Carpenter, MD, MSc; Joseph Colla, MD; Robert W Gibson, PhD; Siera Gollan, PhD, RN; Greg Hall, MD; Lisa Klesges, PhD; Abdullah Kutlar, MD; Matthew Lyon, MD; Cathy L Melvin, PhD, MPH; Sarah Norell, MDes, MFA; Martina Mueller, PhD; Michael B Potter, MD; Rachel Richesson, PhD, MPH, MS; Lynne D Richardson, MD; Gery Ryan, PhD; Lauren Siwney, MD; Marsha Treadwell, PhD; Leslie Zun, MD, MBA; Janelle Armstrong-Brown, PhD; Lisa Cox, MS; Paula Tanabe, PhD

1Washington University School of Medicine, St Louis, MO, United States
2St Jude Children’s Research Hospital, Memphis, TN, United States
3Washington University in St. Louis, St Louis, MO, United States
4RTI International, Durham, NC, United States
5University of Illinois Chicago, Chicago, IL, United States
6Augusta University, Augusta, GA, United States
7Medical University of South Carolina, Charleston, SC, United States
8University of Memphis, Memphis, TN, United States
9University of California San Francisco, San Francisco, CA, United States
10Duke University, Durham, NC, United States
11Icahn School of Medicine at Mount Sinai, New York, NY, United States
12The RAND Corporation, Santa Monica, CA, United States
13Chicago Medical School, North Chicago, IL, United States

Corresponding Author:
Lingzi Luo, MPH, MSW
Washington University School of Medicine
660 S Euclid Ave
St Louis, MO, 63108
United States
Phone: 1 3143323108
Email: lingzi.luo@wustl.edu

Abstract

Background: Individuals living with sickle cell disease often require aggressive treatment of pain associated with vaso-occlusive episodes in the emergency department. Frequently, pain relief is poor. The 2014 National Heart, Lung, and Blood Institute evidence-based guidelines recommended an individualized treatment and monitoring protocol to improve pain management of vaso-occlusive episodes.

Objective: This study will implement an electronic health record–embedded individualized pain plan with provider and patient access in the emergency departments of 8 US academic centers to improve pain treatment for adult patients with sickle cell disease. This study will assess the overall effects of electronic health record–embedded individualized pain plans on improving patient and provider outcomes associated with pain treatment in the emergency department setting and explore barriers and facilitators to the implementation process.

Methods: A preimplementation and postimplementation study is being conducted by all 8 sites that are members of the National Heart, Lung, and Blood Institute–funded Sickle Cell Disease Implementation Consortium. Adults with sickle cell disease aged
18 to 45 years who had a visit to a participating emergency department for vaso-occlusive episodes within 90 days prior to enrollment will be eligible for inclusion. Patients will be enrolled in the clinic or remotely. The target analytical sample size of this study is 160 patient participants (20 per site) who have had an emergency department visit for vaso-occlusive episode treatment at participating emergency departments during the study period. Each site is expected to enroll approximately 40 participants to reach the analytical sample size. The electronic health record–embedded individualized pain plans will be written by the patient’s sickle cell disease provider, and sites will work with the local informatics team to identify the best method to build the electronic health record–embedded individualized pain plan with patient and provider access. Each site will adopt required patient and provider implementation strategies and can choose to adopt optional strategies to improve the uptake and sustainability of the intervention. The study is informed by the Technology Acceptance Model 2 and the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework. Provider and patient baseline survey, follow-up survey within 96 hours of an emergency department vaso-occlusive episode visit, and selected qualitative interviews within 2 weeks of an emergency department visit will be performed to assess the primary outcome, patient-perceived quality of emergency department pain treatment, and additional implementation and intervention outcomes. Electronic health record data will be used to analyze individualized pain plan adherence and additional secondary outcomes, such as hospital admission and readmission rates.

Results: The study is currently enrolling study participants. The active implementation period is 18 months.

Conclusions: This study proposes a structured, framework-informed approach to implement electronic health record–embedded individualized pain plans with both patient and provider access in routine emergency department practice. The results of the study will inform the implementation of electronic health record–embedded individualized pain plans at a larger scale outside of Sickle Cell Disease Implementation Consortium centers.

Trial Registration: ClinicalTrials.gov NCT04584528; https://clinicaltrials.gov/ct2/show/NCT04584528.

International Registered Report Identifier (IRRID): DERR1-10.2196/24818

(JMIR Res Protoc 2021;10(4):e24818) doi:10.2196/24818

KEYWORDS
sickle cell disease; RE-AIM; emergency department care; pain treatment; digital medicine; health innovation; implementation science; patient portal; electronic health record

Introduction

Background
Sickle cell disease is an inherited red blood cell disorder affecting approximately 100,000 people in the United States, predominantly African Americans [1]. In the past few decades, clinical interventions have facilitated significant improvement in patient outcomes. The survival rate to adulthood in sickle cell disease improved from less than 50% in 1970 [2] to nearly 95% in 2010 [3], and the median age at death of patients with sickle cell disease increased from 28 years in 1979 to 43 years in 2014 [4].

Individuals with sickle cell disease often experience acute painful events—vaso-occlusive episodes—during which the transfer of oxygen and nutrients to tissues is decreased because of blood vessel blockage from polymerization [5]. These episodes are characterized by sudden onset of excruciating pain, often requiring high doses of opioids. Historically, pain treatment in the emergency department for individuals with sickle cell disease has been challenging [6]. There is a growing demand to improve the treatment of vaso-occlusive episodes in adults with sickle cell disease, especially in emergency departments, where patients with sickle cell disease require immediate pain treatment [7].

In 2014, the National Heart, Lung, and Blood Institute published an expert panel report [8] of evidence-based recommendations for sickle cell disease management. The use of “an individualized prescribing and monitoring protocol or an SCD-specific protocol whenever possible” [8] in all settings was among the treatment recommendations for vaso-occlusive episode treatment. For pediatric patients with sickle cell disease, studies have found that those treated with individualized pain plans (IPPs) had fewer hospital admissions and readmissions as well as improved pain scores [9-11]. In these studies, providers perceived that having an IPP improved the efficiency and quality of pain management [9,11]. For adult patients with sickle cell disease, however, the literature on the use of IPPs is scarce. A randomized controlled trial has shown that adult patients randomized to patient-specific or weight-based opioid pain plans in the electronic health record (EHR) experienced a significantly greater reduction in pain scores from emergency department arrival to emergency department discharge and a lower hospital admission rate [12]. Another retrospective study found decreased time to first opioid and length of emergency department stay after IPP implementation [13]. EHR use increased from 9.4% in 2008 to 83.8% in 2015 in nonfederal acute care hospitals [14], so these EHR-embedded individualized pain plans are now possible in most hospitals if planned with collaborative efforts and informed by frameworks from Implementation Science [15].

The Sickle Cell Disease Implementation Consortium was established in 2016 to improve the health and well-being of adolescents and young adults with sickle cell disease [16]. The Sickle Cell Disease Implementation Consortium is a cooperative research program of 8 clinical centers; a data coordinating center; and the National Heart, Lung, and Blood Institute to promote quality of care for patients with sickle cell disease between the ages of 15 and 45 years. Investigators in the Sickle Cell Disease Implementation Consortium conducted a systematic
literature review and a comprehensive needs assessment among the 8 participating centers to identify 3 key areas of improvement to address [17]. One of the major opportunities for improvement was the treatment of pain in the adult emergency department.

Patients with sickle cell disease and providers often report frustration with emergency department care, and emergency department providers may have negative attitudes toward individuals living with sickle cell disease [18-21]. These factors may result in patients having lower levels of care satisfaction and the delay or avoidance of emergency department care [17,21,22]. Similarly, from the Sickle Cell Disease Implementation Consortium needs assessment, patients reported being less pleased with their emergency department care compared to their routine care, with only approximately half of participants being satisfied or perceiving adequate quality care in the emergency department [17]. Across 8 sites, 65.7% of the 437 respondents reported that they required emergency department care for acute episodes of pain, and 34.6% of respondents reported 3 or more hospital admissions for pain in the previous year, which is consistent with previous reports in the literature [17]. Slightly fewer than half of the emergency department provider respondents reported that either their emergency department does not have a pain treatment protocol or they were unaware if such a protocol exists [23]. The needs assessment results suggested that establishing standardized treatment in adult emergency departments is key to improved clinical outcomes and care-seeking experiences for adult patients with sickle cell disease across 8 Sickle Cell Disease Implementation Consortium sites. Members from the Sickle Cell Disease Implementation Consortium formed a study workgroup to improve emergency department care for adult patients with sickle cell disease. The workgroup is composed of investigators from the 8 participating Sickle Cell Disease Implementation Consortium sites, representatives from the data coordinating center, and patient, caregiver, and community stakeholders.

Study Aims

Overview

This study has 3 aims involving assessing organizational readiness; the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) of the electronic health record–embedded individualized pain plans (E-IPPs); and barriers and facilitators of E-IPP implementation and use.

Aim 1

The first aim is to assess the overall effectiveness of E-IPPs in improving patient and provider outcomes associated with pain treatment in the adult emergency department setting, with the following subaims: (1) examine the effectiveness of the E-IPP on improving patients’ perceived quality of emergency department pain treatment and (2) examine the effectiveness of the E-IPP on improving providers’ self-efficacy in treating pain for patients with sickle cell disease and perceived quality of emergency department pain treatment.

Aim 2

The second aim is to assess the reach, adoption, implementation, and maintenance of the E-IPP components and implementation strategies at each participating site, with the following subaims: (1) assess the reach of the E-IPP, (2) assess the adoption and implementation of the E-IPP and track implementation strategies adopted by each site, and (3) assess the intent to continue using the E-IPP from a multistakeholder perspective.

Aim 3

The third aim is to assess organizational readiness at the beginning of implementation and barriers and facilitators to the use of E-IPPs.

Methods

Intervention

The intervention for this study is E-IPPs with emergency department provider and patient access. IPPs are records developed by the sickle cell disease providers at each study site based on patients’ outpatient opioid use and analgesic agent normally required for treatment of a vaso-occlusive episode in the emergency department (Textbox 1). Each site will work with its local informatics team to make the E-IPPs available to emergency department providers via the provider EHR interface (Figure 1) and to patients via the EHR patient portal (Figure 2). The IPPs will be reviewed by the patients’ sickle cell disease providers every 6 months. Sites will make adjustments to patients’ IPPs to ensure that all participating patients have all required content and the E-IPPs are easily accessible to emergency department providers. The E-IPPs in the patient portal will allow patients to show their IPPs as a credible source to any emergency department providers in the United States.
Figure 1. An example of an E-IPP emergency department provider interface (Washington University). E-IPP: electronic health record–embedded individualized pain plan.

After clicking on the “Sickle Cell Pain Plan” box, the plan shows up

Figure 2. An example of an E-IPP patient portal (Washington University). E-IPP: electronic health record–embedded individualized pain plan.
Study Setting
This study will take place in all 8 Sickle Cell Disease Implementation Consortium centers (Table 1). Emergency department practices vary. Prior to study implementation, 2 emergency departments had E-IPPs requiring additional content insertion or building plans for some patients, while the remainder either lacked any IPPs or only had IPPs in clinical notes or as a hard copy in a secure server not in the EHR. One of the sites had patient IPP access in the clinical notes before the study, and the rest of the sites did not have patient access to IPPs.

Table 1. Study site characteristics preimplementation.

<table>
<thead>
<tr>
<th>Site</th>
<th>City</th>
<th>Estimated patient population</th>
<th>Community</th>
<th>Setting</th>
<th>Emergency department provider IPPa access</th>
<th>Patient IPP access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augusta University Adult Center for Blood Disorders</td>
<td>Augusta</td>
<td>358</td>
<td>Urban</td>
<td>Academic</td>
<td>IPPs in secured servers in the emergency department for some patients, not in the electronic health record</td>
<td>No</td>
</tr>
<tr>
<td>Duke University Medical Center</td>
<td>Durham</td>
<td>450</td>
<td>Suburban</td>
<td>Academic</td>
<td>Electronic health record–embedded IPPs, but they do not have provider contact</td>
<td>No</td>
</tr>
<tr>
<td>Mount Sinai Hospital</td>
<td>New York</td>
<td>175</td>
<td>Urban</td>
<td>Academic</td>
<td>IPPs in clinical notes</td>
<td>No</td>
</tr>
<tr>
<td>Methodist University Hospital</td>
<td>Memphis</td>
<td>350</td>
<td>Urban</td>
<td>Private hospital</td>
<td>IPPs in clinical notes</td>
<td>IPPs in clinical notes</td>
</tr>
<tr>
<td>Barnes-Jewish Hospital</td>
<td>St. Louis</td>
<td>300</td>
<td>Urban</td>
<td>Academic</td>
<td>No IPPs</td>
<td>No</td>
</tr>
<tr>
<td>University of California at San Francisco Benioff Children’s Hospital Oakland</td>
<td>Oakland</td>
<td>286</td>
<td>Urban</td>
<td>Academic</td>
<td>IPPs in clinical notes</td>
<td>No</td>
</tr>
<tr>
<td>University of Illinois Hospital &amp; Health Sciences System, Sickle Cell Center</td>
<td>Chicago</td>
<td>600</td>
<td>Urban</td>
<td>Academic</td>
<td>IPPs in emergency department and clinical notes for select patients</td>
<td>No</td>
</tr>
<tr>
<td>Medical University of South Carolina Health Emergency Department</td>
<td>Charleston</td>
<td>400</td>
<td>Urban</td>
<td>Academic</td>
<td>Electronic health record–embedded IPPs</td>
<td>No</td>
</tr>
</tbody>
</table>

aIPP: individualized pain plan.

Study Frameworks and Models
This study is guided by the RE-AIM framework, developed in 1999 to assess five dimensions (Reach, Effectiveness, Adoption, Implementation, and Maintenance) [24-26]. Each dimension of RE-AIM will be evaluated with measures at the patient, provider, or organizational and system levels where appropriate. RE-AIM was selected for this intervention because of its focus on dimensions of intervention design and implementation processes that can either facilitate or impede beneficial outcomes and can be replicated and sustained in diverse clinical settings [26,27].

To complement the RE-AIM framework, this study will use a simplified version of the Technology Acceptance Model 2 (TAM2) [28] to understand E-IPP use (Figure 3). Although the TAM2 is a comprehensive model to explain technology acceptance and use, it includes additional constructs that are impractical to be measured in the emergency department setting. We have simplified the TAM2 model and will use its main constructs to help explain the intention and actual use of the E-IPPs.

Figure 3. Simplified Technology Acceptance Model 2.

The study protocol was also developed based on the Standard Protocol Items for Clinical Trials [29], a guideline for the
minimum content of a clinical trial protocol, and the Standards for Reporting Implementation Studies [30], a guideline to improve the reporting of implementation studies.

**Study Design**

The study uses a preimplementation and postimplementation design in which the IPPs will be embedded in patient EHRs, with both provider and patient having access to view the E-IPPs. Multimedia Appendix 1 shows the program logic model specifying the determinants, implementation strategies, mechanisms of action, and outcomes [31]. The organization of determinants is informed by the Consolidated Framework of Implementation Research, which provides a comprehensive list of constructs that would influence implementation [32], and the study’s readiness assessment, described in the following section. Figure 4 shows the study flow for each site.

Given that sickle cell disease is a rare disease and most patients with sickle cell disease do not have frequent emergency department visits per year, the preimplementation and postimplementation design without control groups is the most feasible study design. Sites will overenroll participants to ensure that enough participants will have a qualified emergency department visit during the implementation period. Additionally, it took a few months to 2 years for sites to work with their local informatics teams to make the E-IPPs available to emergency department providers and patients; therefore, a waitlist-control or stepped-wedged design is not feasible given the study timeline. The study rollout will allow for emergency department provider training of the E-IPP, development and implementation of the E-IPPs, and evaluation postimplementation. Because of the COVID-19 pandemic, sites have adjusted to the rollout timeline. Two sites began the project in September 2020. The active implementation period is 12 months from the first day of patient enrollment, and the overall implementation phase will be 18 months.

**Implementation Strategies**

Based on existing evidence, workgroup expertise, and the feasibility of implementing the E-IPP, the workgroup identified required and optional implementation strategies at the patient and provider levels to facilitate implementation and use of the E-IPP (Table 2).
Table 2. Implementation strategies at the patient and provider levels.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Patient</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required</td>
<td>• Study orientation&lt;br&gt;• Download and install electronic health record patient portal app&lt;br&gt;• Video demonstration of how to access the E-IPP&lt;sup&gt;a&lt;/sup&gt;&lt;br&gt;• Patients will be asked to show the staff how to access their pain plans via their phone (a teach-back method)&lt;br&gt;• Provide a wallet card with instructions to access the E-IPP and staff contact information to take home&lt;br&gt;• Quarterly reminder text with video demonstration</td>
<td>• Provider training with a tracking log&lt;br&gt;• A 5-minute video addressing stigma and the actual prevalence of opioid addiction in sickle cell disease&lt;sup&gt;[33]&lt;/sup&gt;&lt;br&gt;• Introduction to the study and demonstration of how to access the E-IPP</td>
</tr>
<tr>
<td>Optional</td>
<td>• Communication script: provide a script at orientation for patients to practice how to communicate with emergency department providers about the E-IPP</td>
<td>• Electronic health record–embedded prompts to remind emergency department triage clinicians of E-IPP&lt;br&gt;• Quarterly booster education sessions&lt;br&gt;• Disseminate and educate through podcasts, blogs, or journal clubs&lt;sup&gt;[34]&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>E-IPP: electronic health record–embedded individualized pain plan.

Participants and Eligibility Criteria

Site coordinators will keep a tracking log for patient and provider participants (Table 3).

Table 3. Eligibility criteria.

<table>
<thead>
<tr>
<th>Entity</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participating sites</td>
<td></td>
</tr>
<tr>
<td>Inclusion</td>
<td>Site hematologist or sickle cell disease provider willing to write an IPP&lt;sup&gt;a&lt;/sup&gt; for patients meeting eligibility criteria&lt;br&gt;Informatics resources available to support all aspects of the intervention&lt;br&gt;Placement of the E-IPP&lt;sup&gt;b&lt;/sup&gt; that is accessible to both the provider and patient&lt;br&gt;Ability to support text messaging to patients for survey administration&lt;br&gt;Support from emergency medicine and nursing leadership to agree to follow the IPP unless there is a contraindication at the time of the emergency department vaso-occlusive episode visit</td>
</tr>
<tr>
<td>Exclusion</td>
<td>None</td>
</tr>
<tr>
<td>Patient participants</td>
<td></td>
</tr>
<tr>
<td>Inclusion</td>
<td>Confirmed sickle cell disease diagnosis, defined as supported by documentation in the medical record of a positive test for 1 of the following genotypes: Hb SS, Hb SC, Hb Sβ-thalassemia, Hb SO, Hb SD, Hb SG, Hb SE, or Hb SF. If no medical record is available, the enrolling site will conduct its own laboratory test as confirmation&lt;br&gt;English speaking&lt;br&gt;Age 18–45 years (National Heart, Lung, and Blood Institute grant requirement)&lt;br&gt;Access to either an Android or iOS cellular or mobile smartphone, with access to text messaging and internet&lt;br&gt;At least 1 vaso-occlusive episode visit to the participating site’s emergency department within the past 90 days prior to enrollment&lt;br&gt;At least 1 visit at the study site sickle cell disease clinic within the past 12 months&lt;br&gt;Willing and cognitively able to give informed consent</td>
</tr>
<tr>
<td>Exclusion</td>
<td>Site hematologist or sickle cell disease provider states patient should not have a protocol or should not be administered opioids</td>
</tr>
<tr>
<td>Provider participants</td>
<td></td>
</tr>
<tr>
<td>Inclusion</td>
<td>All emergency physicians, nurses, physician assistants, and nurse practitioners who work in the study emergency department will have access to the E-IPPs as routine practice and physicians are asked to complete baseline surveys at enrollment for research purposes</td>
</tr>
<tr>
<td>Exclusion</td>
<td>None</td>
</tr>
</tbody>
</table>

<sup>a</sup>IPP: individualized pain plan.

<sup>b</sup>E-IPP: electronic health record–embedded individualized pain plan.
Recruitment

Patient Recruitment and Procedures

Potential participants will be approached in person, by phone, or via electronic media about enrolling in the study. All patients enrolled in the Sickle Cell Disease Implementation Consortium registry and screened as eligible will be contacted by research staff for participation in the emergency department project. Additional participants meeting criteria at participating Sickle Cell Disease Implementation Consortium sites not previously enrolled in the registry are also eligible and may be recruited in the clinic, during hospitalizations, or at community events by research staff. Patient participants will provide informed consent before study participation.

Provider Recruitment

All emergency department providers will be able to access the E-IPPs and will receive training on the protocol at staff meetings, resident conferences, or through emails. All emergency department providers will be invited to complete a baseline survey at enrollment with an information sheet introducing the study. If an enrolled patient participant has a qualified emergency department visit and the corresponding emergency department provider has not received training or completed the baseline survey, the study team will send a provider follow-up survey to the provider at enrollment with an information sheet introducing the study.

A qualified emergency department visit is defined as (1) a visit for an enrolled patient to a participating emergency department, the reason for the visit is a vaso-occlusive episode, and the patient has an E-IPP at the time of the emergency department visit; and (2) the patient’s first emergency department vaso-occlusive episode visit of the month, after enrollment, within the 12 month study period for that patient.

Data Collection

The data collected in this study will consist of a readiness assessment; EHR data retrieval from patient records; patient baseline and follow-up surveys; provider baseline and follow-up surveys; and patient, provider, and emergency department administrator interviews. All measures (other than the readiness assessment) are matched with the RE-AIM framework (Multimedia Appendix 2).

Readiness Assessment

The participating centers will administer a Readiness Diagnostic Scale (RDS) developed by members of the Consortium’s Implementation Research Committee at the beginning of the implementation period [35]. The RDS is a quantitative assessment to capture 3 interrelated dimensions of organizational readiness: general capacities (34 items), assessing participating emergency department’s existing emergency department practice and how it functions overall; intervention-specific capacities (13 items), assessing participating emergency department capacity to use the E-IPP; and motivation (17 items), assessing how well the emergency department facilitates physicians’ willingness to use the E-IPP.

Each site is required to have at least 5 total completed RDSs from a study team researcher, emergency department administrator, emergency department physician, and emergency department nurse to capture different system-level roles in clinical care.

EHR Data Retrieval

Each participating site will perform EHR data retrieval for enrolled patients’ past 12 months ED visits and qualified ED visits to collect: hospital admission rate; 7- and 30-day emergency department revisit rate; 7- and 30-day hospital readmission rate; time to first dose of analgesic agent; and administered analgesic agent, dose, and route. At the end of the project, the site will also retrieve the following from the EHR: number of E-IPPs written and time of writing or updating, number of new E-IPPs available in the EHR and time of becoming available, and number of new patients at each site who are being offered access to their pain plan in the EHR after study enrollment targets are met. Data retrieval will follow a manual of procedures created by the workgroup to reduce bias [36,37].

Patient Baseline and Follow-up Survey

Each participating site will administer a brief patient baseline survey at the time of patient enrollment that will assess patient demographic data, patient-perceived quality of emergency department pain treatment for the last emergency department visit (within 90 days of enrollment), and how well the patient knows how to use the patient portal. Within 96 hours of a qualified emergency department visit, the research team will send a follow-up survey by text message to the patient that will assess patient’s perceived quality of emergency department pain treatment for this visit, how well they know how to use the patient portal, perceived ease of use of the E-IPP, patient and emergency department provider use of E-IPP during the last visit, satisfaction with the E-IPP, and intent to use the E-IPP for next emergency department vaso-occlusive episode visit. Patients will receive up to 3 follow-up surveys during the study period.

Patient Interview

After a center has had 5 patients with qualified emergency department visits, its team will begin to invite patients for interviews within 2 weeks of a qualified emergency department visit. The patient interview will assess the patient’s experiences using (or not using) the E-IPP, what was helpful and challenging about using the E-IPP, if pain treatment has changed (or not changed) because of the E-IPP, patient satisfaction with the E-IPP, proposed recommendations to improve the E-IPP, intent to use the E-IPP in the future, and additional questions on implementation strategies such as the wallet card. The research team will use purposeful sampling for qualitative data collection [38], based on matched patient and provider survey responses and implementation outcomes of the IPPs.

Provider Baseline and Follow-up Survey

Each participating site will administer a brief provider baseline survey at the time of provider enrollment that will assess provider-perceived quality of emergency department pain treatment and provider’s self-efficacy to manage acute pain episodes for patients with sickle cell disease. Within 96 hours of a qualified emergency department visit, the research team
will send a follow-up survey via email to the enrolled provider who ordered the first analgesic. The provider follow-up survey will assess use of E-IPP, ease of E-IPP use, E-IPP adherence, perceived quality of emergency department pain treatment, satisfaction with the E-IPP, and intent to use the E-IPP in the future.

**Provider Interview**

After a center has had 5 patients with qualified emergency department visits, the team will begin to invite providers who ordered the first analgesic for interviews within 2 weeks of a qualified emergency department visit. The provider interview will assess provider’s experiences using (or not using) the E-IPPs, what was helpful and challenging about using the E-IPPs, proposed recommendations to improve the E-IPPs, intent to use the E-IPP in the future, and additional questions on implementation strategies. The research team will use purposeful sampling for qualitative data collection [38], based on matched patient and provider survey responses and implementation outcomes of the E-IPPs.

**Emergency Department Administrator Interview**

At the end of the implementation period, each site team will invite emergency department administrators to a postimplementation interview that will assess barriers and facilitators to intervention implementation, emergency department administrator’s experiences implementing and using the E-IPP, and emergency department administrator’s intent to continue using the E-IPP in the future.

**Tracking and Reporting Implementation Strategies**

A parallel study has been funded to capture the planned and actual implementation strategies employed by each site [39]. Semistructured interviews with 3 stakeholders from each site before participant recruitment and at the end of the intervention, as well as quantitative surveys at the midpoint of project implementation, will capture the barriers and facilitators of the strategies planned and whether there have been any adaptations made. Questions have been informed by the implementation and maintenance domains of the RE-AIM framework.

**Statistical Analysis Power and Sample Size**

The study plans to have an analytical sample size of 160 patient participants, 20 for each site, for its primary outcome analysis. Each site is expected to enroll approximately 40 participants to reach the analytical sample size.

Sample size was calculated for the primary outcome, patient satisfaction with emergency department pain treatment through their perceived quality of emergency department pain treatment. For the power calculations, we used results from the Sickle Cell Disease Implementation Consortium needs assessment [17, 23], which used similar measurements to assess patient satisfaction with emergency department pain treatment. We assumed that the correlations between responses would not change and conducted a simulation to determine statistical power to detect a treatment difference of half a standard deviation. The sample estimation model has several assumptions: the intervention effect will vary across sites with the standard deviation for the random intervention effect set at 0.447; within-site correlation (commonly referred to as intercluster correlation) of baseline measurements, calculated as the ratio of site-level variance over the total variance, is set at 0.10; and within-participant correlation, calculated as the ratio of individual-level variance over total individual variance, is set at 0.50, reflecting the expectation that the preintervention composite score is predictive of the postintervention score.

Based on the Sickle Cell Disease Implementation Consortium registry data, we estimated that at least 50% of the study patient participants will have at least 1 qualified emergency department visit during the study period and that approximately 20% of the participants will be lost to follow-up and provide no posttest data. The total number of patients a site will recruit is approximately 40. The results indicate that the study will have >90% power to reject the null hypothesis when the intervention results in an average improvement in patient-perceived quality of emergency department care of 0.5 standard deviations under the assumed conditions.

**Data Analysis**

**Primary Outcome**

The primary outcome is patient satisfaction with emergency department pain treatment through their perceived quality of emergency department pain treatment (Multimedia Appendix 2). This is measured with a composite of 3 questions developed based on the Adult Sickle Cell Quality of Life Measure Quality of Care measure [6]. The impact of the intervention on the primary outcome will be estimated using a linear mixed model. A 3-level model with random effects for site, participants nested within site, and site-to-site variation in the treatment effect and a fixed effect for treatment response will be specified to assess the impact of the intervention. The primary analysis will be based on the preintervention composite score and the score at 1 emergency department visit per participant, using the first visit for those with more than one. The model can accommodate multiple visits per participant by adding a covariance matrix to account for repeated observations.

**Other Outcomes**

The secondary outcomes of hospital admission rate (a count outcome, as multiple hospital admissions are possible), 7-day emergency department vaso-occlusive episode revisit (a dichotomous outcome), and time to first dose, and other provider-level outcomes will be analyzed with generalized linear mixed models, where the outcome will be specified as a Poisson or binomial variable with a logit function as needed by the data type of outcome. The generalized linear mixed models would follow the same format as the linear mixed model with fixed effects for time, intervention, and covariates and random effects for participants within sites. The secondary outcome of satisfaction with IPP, measured with a single Likert scale question for both patient and provider, will be analyzed as a continuous outcome with a linear mixed model as detailed above for the primary outcome. In all cases, the analysis will estimate an intervention effect that will compare the secondary outcome at baseline to after the intervention is delivered. Power calculations indicate that the study sample will be too small to detect statistically significant differences in these noncontiguous
variables, but estimates will be provided to inform future research.

The analysis of the quantitative readiness assessment result and the reach, adoption, implementation, and maintenance outcomes will be descriptive. As a preliminary attempt, at the end of the intervention, the results of the readiness assessment will inform our interpretation of the implementation and effectiveness outcomes. Sites with a lower level of readiness may face challenges during implementation, which will affect their implementation outcome and intervention effectiveness.

For qualitative interviews, the workgroup members will develop a codebook informed by the RE-AIM framework for data analysis. Using a deductive approach, the codebook will include an initial list of codes to be used in the analysis and definitions or operationalized examples for each code. Analysts will revise the codebook as necessary to hone definitions to increase consistency in coding across the research teams. Data will be compiled into different stakeholder groups by themes and analyzed across the study sites.

**Results**

At the time of submission, this study was approved by the institutional review board of Duke University (IRB Pro 000073506), Icahn School of Medicine at Mount Sinai (IRB 20-04302), Medical University of South Carolina (IRB Pro 00097830), Washington University School of Medicine (IRB 202006209), University of California San Francisco (IRB 20-30981), University of Illinois Chicago (IRB 2016-02998). The full protocol and study manual of procedures are available from the corresponding author upon request. The study is registered (NCT04584528) and enrolling participants. Sites that did not have IPPs for all patients are working with the sickle cell disease providers to build the plans with patients during clinic visits. Because of the COVID-19 outbreak, sites have changed the planned recruitment timeline and will continue enrolling until the target is reached.

**Discussion**

With the rapid expansion of the use of EHR in the past decade, building an EHR-embedded treatment plan for different patient populations has gained increased interest, especially in the emergency department, where decision making is time sensitive. Previous research has explored various decision support tools, including treatment plans, to facilitate emergency department provider decision making [11,40-42]. The E-IPP intervention is a direct result of a large-scale needs assessment effort and is designed to meet the needs of participating sites and patient populations. Guided by the RE-AIM theoretical framework and a widely used TAM2 model explaining technology uptake, the protocol proposes to test the effectiveness of the E-IPP while also describing how the intervention is implemented and assessing implementation outcomes in detail.

Previous research on individualized treatment plans for patients with sickle cell disease has largely been conducted in pediatric settings [9-11]. These studies demonstrated that providers perceived that digitization of the treatment plans could help improve the efficiency and quality of pain management [9-11]. So far, only 1 randomized controlled trial [12] has tested the effectiveness of an individualized treatment plan compared to a weight-based plan for adult patients with sickle cell disease and showed significant efficacy, but the results were not generalizable, as it was conducted in only 2 emergency departments with 52 patients. While previous studies have focused on assessing patient outcomes, this study will assess multilevel outcomes, including patient experience and satisfaction. With a comprehensive understanding of the uptake and implementation of the E-IPPs, the intervention will be more likely to successfully be disseminated to other emergency departments on a large scale.

The most novel component of this protocol is making the E-IPP available to the patient. Many sickle cell centers already make IPPs available to the emergency department provider, but few have been able to make these accessible to the patient. Among 8 Sickle Cell Disease Implementation Consortium sites, only 1 site had patient IPP access through the clinical notes. With this project, patients will have easy E-IPP access and will be able to present their IPPs to the emergency department provider directly via the patient portal app on their cellular device. In the past decade, smartphone ownership has increased significantly in the United States, even among low- and moderate-income communities [43]. Over 70% of individuals with an annual income of less than $30,000 reported smartphone ownership in 2019, and the rates of smartphone ownership among patients with sickle cell disease are similar [43-45]. Previous studies [46-48] have identified barriers to patients accessing and using an EHR, and participants with limited health literacy may need additional time to navigate the EHR patient portal. Previous literature has demonstrated substantial disparities in portal use, indicating that vulnerable populations, such as racial and ethnic minorities and individuals with low socioeconomic status, are less likely to use patient portals, which is relevant for sickle cell disease, as the majority of individuals living with sickle cell disease are African American [49,50]. This study will accommodate many barriers by providing a one-on-one session to help patients install the patient portal for E-IPP access and by providing a wallet card with instructions and videos of how to access their E-IPPs. As previous exploratory research on disease-specific patient portals in patients with sickle cell disease has achieved high acceptability, the E-IPPs, within a few clicks from the main menu, are likely to be viewed favorably by study participants [50]. The study will be able to explore whether patient access to E-IPPs will help with their treatment experiences in the emergency departments.

For the protocol, a significant challenge is to balance the desire for robust study design with practical needs in the emergency departments, which is common for implementation research studies [51]. Acknowledging that a preimplementation and postimplementation study design without control groups has limitations, it is the only feasible solution given that sickle cell disease is a rare disease and study sites must overenroll participants to collect data on enough emergency department visits to generate meaningful results. All emergency department sites want to adopt E-IPP implementation to improve practice as soon as possible. The protocol uses comprehensive mixed
methods data collection [52] to compensate for the study design to achieve the 3 study aims. Some of the activities will be exploratory in nature and provide preliminary data to inform future work expanding the E-IPP on a larger scale and beyond academic settings.

Another significant challenge is that the Sickle Cell Disease Implementation Consortium consists of 8 academic sites, and local practices vary. Because the EHR systems may be structured in different ways, sites have spent an extensive amount of time working with their local informatics teams to make the E-IPPs available in both patient and provider channels. The protocol has adopted several measures to capture and address this diversity, including surveying sites about their existing practices, incorporating the readiness assessment component to help sites understand capacity and motivations that may be unique to their site, and a separate protocol to solely focus on implementation strategies and adaptations at the site level.

In summary, this study proposes a framework-informed, structured approach to implement a guideline recommendation in routine emergency department practice. It is the first multisite study investigating the effectiveness of E-IPPs in adult emergency departments at both the provider and patient levels. The results of the study will inform the implementation of IPPs with emergency department provider and patient access at a larger scale.

Acknowledgments
The Sickle Cell Disease Implementation Consortium has been supported by US Federal Government cooperative agreements (HL133948, HL133964, HL133990, HL133996, HL133994, HL133997, HL134004, HL134007, and HL134042) from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities in Bethesda, Maryland. We thank our Sickle Cell Disease Implementation Consortium stakeholders for their contributions to the project: Kamilah Bailey, Johnny Gamble, Georgia Boateng, Leroy Ediage, Donnette Carroll, Danielle Jamison, Justin Flowers, Amanda Fuller, Sorivette Surillo, Beverly Chukwadozie, and Rosemary Britts. A full list of Sickle Cell Disease Implementation Consortium members can be found in Multimedia Appendix 3.

Conflicts of Interest
PT has received funding from the National Heart, Lung, and Blood Institute and is a consultant for CSL Behring LLC.

References


Abbreviations
- EHR: electronic health record
- E-IPP: electronic health record–embedded individualized pain plan
- IPP: individualized pain plan
- RDS: Readiness Diagnostic Scale
- RE-AIM: Reach, Efficacy, Adoption, Implementation, and Maintenance Model
- TAM2: Technology Acceptance Model 2

Edited by T Derrick; submitted 15.12.20; this is a non–peer-reviewed article; accepted 18.12.20; published 16.04.21.

Please cite as:
URL: https://www.researchprotocols.org/2021/4/e24818
doi:10.2196/24818
PMID:33861209

©Lingzi Luo, Allison A King, Yvonne Carroll, Ana A Baumann, Donald Brambilla, Christopher R Carpenter, Joseph Colla, Robert W Gibson, Siera Gollan, Greg Hall, Lisa Klesges, Abdullah Kutlar, Matthew Lyon, Cathy L Melvin, Sarah Norell, Martina Mueller, Michael B Potter, Rachel Richesson, Lynne D Richardson, Gery Ryan, Lauren Siewny, Marsha Treadwell, Leslie Zun, Janelle Armstrong-Brown, Lisa Cox, Paula Tanabe. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 16.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original work on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Digital Epidemiologic Research on Multilevel Risks for HIV Acquisition and Other Health Outcomes Among Transgender Women in Eastern and Southern United States: Protocol for an Online Cohort

Andrea L Wirtz¹, PhD; Erin E Cooney¹, MSPH; Megan Stevenson¹, MSPH; Asa Radix², MD, PhD; Tonia Poteat³, PhD, PA-C; Andrew J Wawrzyniak⁴, PhD; Christopher M Cannon⁵, MPH; Jason S Schneider⁶, MD; J Sonya Haw⁷, MD; James Case⁷, MBA; Keri N Althoff⁴, PhD; Elizabeth Humes¹, MHS; Kenneth H Mayer⁸,⁹, MD; Chris Beyrer¹, MD; Allan E Rodriguez¹⁰, MD; Sari L Reisner⁹,¹¹,¹²,¹³, SCD; American Cohort To Study HIV Acquisition Among Transgender Women (LITE) Study Group¹

¹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States
²Callen-Lorde Community Health Center, New York, NY, United States
³Center for Health Equity Research, University of North Carolina School of Medicine, Chapel Hill, NC, United States
⁴Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, FL, United States
⁵Research and Evaluation, Whitman-Walker Institute, Washington DC, DC, United States
⁶Emory University School of Medicine, Atlanta, GA, United States
⁷School of Nursing, Johns Hopkins University, Baltimore, MD, United States
⁸Beth Israel Deaconess Medical Center, Harvard Medical School, Harvard University, Boston, MA, United States
⁹The Fenway Institute, Fenway Health, Boston, MA, United States
¹⁰Division of Infectious Diseases, Department of Medicine, University of Miami Miller School of Medicine, Miami, FL, United States
¹¹Division of Endocrinology, Diabetes, and Hypertension, Brigham Women’s Hospital, Boston, MA, United States
¹²Department of Medicine, Harvard Medical School, Harvard University, Boston, MA, United States
¹³Department of Epidemiology, Harvard TH Chan School of Public Health, Harvard University, Boston, MA, United States

Corresponding Author:
Andrea L Wirtz, PhD
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
Johns Hopkins University
615 N Wolfe St
W3505
Baltimore, MD, 21205-4803
United States
Phone: 1 4105020800
Email: awirtz1@jhu.edu

Abstract

Background: The HIV epidemic disproportionately impacts transgender women in the United States. Cohort studies identify unique risks for affected populations, but use of facility-based methods may bias findings towards individuals living in research catchment areas, more engaged in health services, or, in the case of transgender populations, those who are open about their transgender identity. Digital clinical trials and other online research methods are increasingly common, providing opportunity to reach those not commonly engaged in research. Simultaneously, there is a need to understand potential biases associated with digital research, how these methods perform, and whether they are accepted across populations.

Objective: This study aims to assess the feasibility of developing and implementing an online cohort of transgender women to assess risks for HIV acquisition and other health experiences. Further, this study aims to evaluate how an online cohort compares to a site-based, technology-enhanced cohort for epidemiologic research. The overarching goal is to estimate incidence of HIV and other health outcomes among transgender women in eastern and southern United States.
Methods: This substudy is part of a larger multisite prospective cohort (LITE) conducted among transgender women, which also includes a site-based, technology-enhanced cohort in 6 eastern and southern US cities. The online cohort was launched to enroll and follow participants across 72 cities in the same region and with similar demographic characteristics as the site-based cohort. Participants are followed for 24 months. Adult transgender women are recruited via convenience sampling (eg, peer referrals, social media, and dating apps). Participants reporting negative or unknown HIV status are enrolled in a baseline study visit, complete a sociobehavioral survey, and provide oral fluid specimens to test for HIV. Participants not living with HIV (lab-confirmed) at baseline are offered enrollment into the cohort; follow-up assessments occur every 6 months.

Results: Enrollment into the online cohort launched in January 2019. Active recruitment stopped in May 2019, and enrollment officially closed in August 2020. A total of 580 participants enrolled into and are followed in the cohort. A recruitment-enrollment cascade was observed across screening, consent, and completion of study activities. Implementation experiences with HIV test kits highlight the need for heavy staff engagement to support participant engagement, visit completion, and retention, even with automated digital procedures.

Conclusions: This study is responsive to increasing research interest in digital observational and intervention research, particularly for populations who are most affected by the HIV epidemic and for those who may otherwise not participate in person. The progression across stages of the recruitment-enrollment cascade provides useful insight for implementation of cohort studies in the online environment.

International Registered Report Identifier (IRRID): DERR1-10.2196/29152


KEYWORDS
transgender persons; United States; cohort studies; digital research; HIV infection; HIV testing; public health; online health; transgender; HIV

Introduction

Background

Transgender women are one of the most affected populations with respect to the HIV epidemic in the United States and are prioritized in the US strategy to end the HIV epidemic [1,2]. Multiple biological, behavioral, and social risks for HIV infection among transgender women are driven by, or are concomitant with, structural barriers that limit access to HIV prevention, testing, care, and health services [3]. Across the United States, transgender women report a high prevalence of sexual assault and violence, homelessness, unemployment, substance use, and low health insurance coverage [4-12]. These factors increase the likelihood of condomless sex, thereby increasing HIV acquisition risk. Due to these multilevel vulnerabilities, transgender women in the United States experience an estimated 14% laboratory-confirmed HIV prevalence (meta-analysis) — more than 42 times the national HIV prevalence in the United States of 0.3% [13]. Significant disparities exist across racial groups and ethnicity: HIV prevalence is estimated to be 44% among Black transgender women and 25% among Latina transgender women [13]. Gender-based discrimination, stigma in health care, and prioritization of basic needs, including gender-affirming care, impede access and uptake of HIV testing, prevention, care, and treatment [12,14,15]. Effective and acceptable HIV prevention interventions tailored to transgender women are urgently needed and require population-specific insight gained from observational and qualitative research [16-21].

HIV prevention research among transgender women faces several challenges. The lack of transgender-specific marketing and misgendering during research (referring to a person by an incorrect name or pronoun) have left a legacy of wariness about HIV research [22]. Inclusion of transgender women as a subset of a larger study population has produced small samples of transgender women, limited scientific inference, and provided insufficient information about acceptable and effective HIV prevention interventions. Despite prior participation in HIV research, transgender women participants have reported that these efforts rarely benefit their communities [22], highlighting the importance of population-specific and community-engaged research to identify and respond to community needs.

In March 2018, we launched the first multisite cohort to assess HIV incidence among transgender women in the United States (known to participants as the LITE study; NIH UG3/UH3AI133669). This cohort has enabled the development and refinement of technology-enhanced methods to reach and retain transgender women in site-based HIV research. The cohort includes 6 technology-enhanced, site-based cohorts in Boston, New York City, Baltimore, Washington DC, Atlanta, and Miami, providing a combination of facility-based in-person and remote assessments for participants over at least 24 months of follow-up [23]. Facility-based visits include self-administered surveys along with HIV and sexually transmitted infection (STI) testing. During remote visits, participants utilize a hybrid app to complete surveys and HIV self-testing, report the results, and upload photos of the test results for validation. This method allowed us to enroll and retain 731 transgender women with varying levels of consistent technology access across these 6 cities.

Cohort studies identify unique risks for affected populations, but the traditional use of facility-based methods may bias findings towards individuals living in research catchment areas, more engaged in health services, or, in the case of transgender populations, those who are open about their transgender identity. Novel, technology-based methods are rapidly increasing in...
the unique health needs of transgender women. Other health interventions that are developed explicitly to meet the course of follow-up [23]. The overarching goal of these HIV-uninfected participants and the HIV Care Continuum (4) estimate the HIV Prevention Continuum (HIVPC) among high-risk eastern and southern US areas, trends in incidence, transgender women in the eastern and southern United States; HIV-uninfected transgender women into a prospective cohort; which are to (1) determine the efficiency and acceptability of implementation of an online cohort compared to a site-based, exclusively on online or site-based cohort methods, limiting the ability to answer these questions.

Objectives
The objective of the online cohort is to assess the feasibility and qualitative differences in the development and implementation of an online cohort compared to a site-based, technology-enhanced cohort for HIV research. This online cohort will also contribute to the broader aims of the LITE study, which are to (1) determine the efficiency and acceptability of novel, technology-infused recruitment methods to enroll HIV-uninfected transgender women into a prospective cohort; (2) describe the demographic, socioeconomic, behavioral, and physical and mental health profiles of HIV-uninfected transgender women in the eastern and southern United States; (3) estimate HIV incidence among transgender women in high-risk eastern and southern US areas, trends in incidence, and associated individual, social, and structural risk factors; and (4) estimate the HIV Prevention Continuum (HIVPC) among HIV-uninfected participants and the HIV Care Continuum (HIVCC) among transgender women who acquire HIV over the course of follow-up [23]. The overarching goal of these cohorts is to inform subsequent HIV prevention, HIV care, and other health interventions that are developed explicitly to meet the unique health needs of transgender women.

Methods
Design
This protocol describes the methods of a “sister” online cohort for transgender women in eastern and southern United States. Online cohort methods are largely developed to match the site-based methods conducted under the LITE study, to the closest extent possible, which has previously been described [23]. The protocol described here provides an overview of the online cohort methods, highlighting distinctions between the online and site-based cohorts. Notably, the online cohort utilizes strictly remote study assessments that include self-administered, electronic surveys and the use of self-collected, oral fluid specimens, which are shipped to a central laboratory for HIV-1 testing. Study assessments occur every 6 months, rather than quarterly visits as in the site-based cohorts. STI testing is not performed due to resource constraints.

All cohort methods have been informed by formative qualitative research, which focused on assessing community perspectives of technology-enhanced and online research methods, HIV testing methods, concerns and suggestions related to HIV research among transgender women, and optimal ways to engage the community [22,47,48]. A virtual community advisory board (CAB) facilitates this research by serving as a mechanism for community consultation for both the site-based and online cohorts and is engaged throughout every phase of the research activities. CAB meetings are held every 4-6 months to discuss research activities, concerns and protection of community members, results, and plans for dissemination to the wider community [49,50].

 Sample
The online cohort enrolled and follows adult transgender women who are not living with HIV in 72 eastern and southern US metropolitan areas with population sizes >100,000. Participating cities for the online cohort were selected on the basis of similar demographic characteristics to the initial 6 cities (Boston, New York City, Baltimore, Washington DC, Atlanta, and Miami) where the site-based cohort is implemented.

Preliminary, self-administered screening assesses for the following inclusion criteria: aged 18 years or older; no prior diagnosis of HIV infection; speaks and understands English or Spanish; and residence in or within a 90-mile radius of one of the following cities: Atlanta, GA; Baltimore, MD; Baton Rouge, LA; Birmingham, AL; Boston, MA; Charlotte, NC; Chicago, IL; Columbia, SC; Detroit, MI; Jackson, MS; Jacksonville, FL; Memphis, TN; Miami, FL; Nashville, TN; New Haven, CT; New Orleans, LA; New York, NY; Newark, NJ; Orlando, FL; Philadelphia, PA; Pittsburgh, PA; Providence, RI; Tampa, FL; Washington, DC; Worcester, MA; Allentown, PA; Aurora, IL; Bridgeport, CT; Buffalo City, NY; Charleston, SC; Charleston, NC; Chattanooga, TN; Cincinnati, OH; Clarksville, TN; Cleveland, OH; Columbus, GA; Columbus, OH; Fayetteville, NC; Flint, MI; Fort Wayne, IN; Ft. Lauderdale, FL; Hartford, CT; Hollywood, FL; Huntsville, AL; Indianapolis, IN; Jersey City, NJ; Knoxville, TN; Lafayette, LA; Lansing, MI; Lexington, KY; Louisville, KY; Mobile, AL; Montgomery, AL;
Norfolk, VA; Palm Bay, FL; Port St. Lucie, FL; Raleigh/Durham, NC; Richmond, VA; Rochester City, NY; Rockford, IL; Savannah, GA; Shreveport, LA; Springfield, NC; Stamford, CT; Syracuse, NY; Tallahassee, FL; Tampa/Clearwater/St. Petersburg, FL; Toledo, OH; Virginia Beach/ Newport News, VA; Waterbury, CT; Wilmington, NC; and Yonkers City, NY. Participants’ transgender status was determined during screening using the recommended two-step method (assigned male sex at birth and current gender identity as a woman or along the trans feminine spectrum) [51,52], inclusive of binary and nonbinary people. Participants who were eligible based on preliminary screening and provided electronic consent to participate were then asked to complete the baseline survey and are sent an oral fluid specimen collection kit for laboratory verification of HIV status and final determination of eligibility in the cohort.

Prior to the consent and enrollment process, participants were asked to complete a 2-factor authentication process (to ensure they are not duplicate enrollees or robotic enrollments). Individuals who are duplicate enrollments or who report concurrent enrollment in an HIV prevention trial were ineligible for participation.

**Recruitment and Enrollment**

Recruitment utilized a mix of technology-infused methods, including paid advertisements on Google Ads; social media venues frequented by transgender women including Facebook, Instagram, and Reddit; as well as geosocial networking apps used for dating (eg, Grindr and Black Gay Chat; Figure 1). Online advertisements for these sites permitted targeting of ads to the specified cities and surrounding areas, user profiles, or user search terms based on internal algorithms (no individual-level identifying data are shared with researchers). Our study website synchronized with Google Analytics to permit real-time analysis of geographic locations of users, keywords, and sites that referred them, which allowed us to monitor and adjust recruitment strategies as enrollment proceeded.

**Figure 1.** Image of a social media advertisement.

Advertisements provided a study phone number and link to the study webpage [53]. Participants who followed the link to the webpage could view study information in English or Spanish and access screening and enrollment for the online cohort. Screening assessments were automated with staff review. Individuals living in the 6 site-based cohort cities had the option to enroll in either cohort and could switch cohort at any point over the course of follow-up, depending on their location of residence.

Once a participant provided a unique email on the screening form, was determined to be eligible for the study, and consented to join the study, the system sent a unique link via email. Once this link had been accessed, a verification code was sent via text. On receipt of the text, the participant entered that code into the webpage to complete enrollment. The initial email message that went to consented participants included a link back to this validation page for later use, if a user was not able to provide a phone number at that time. If the code was validated, the user was directed to complete registration and the baseline survey.

Collected identifiers (name, mailing address, telephone number, and email address) are stored separately from the participant survey, and HIV testing data are retained until study end to prevent duplicate enrollments. Individuals using Voice Over Internet Protocol phone numbers are also prevented from enrolling to mitigate fraudulent behaviors or enrollment. An internal algorithm ran real-time checks to identify any duplicate or partially matched identifiers. Identical matches were
prohibited from registering; partial matches were reviewed by
study staff for final determination.

**Data Collection**

Data collection, both survey completion and specimen
collection, are supported through the use of a hybrid app
developed for this study. The hybrid app enables download of
the secure app on Android and iOS operating systems but also
provides an optional web-based platform for non-smartphone
users. We elected to use a hybrid app, recognizing that while
transgender women have high rates of internet use [54], reliance
solely on app-based data collection may threaten study retention.

**Hybrid App**

The app, which was developed for the site-based cohorts, was
modified for the online cohort to display the participants’ study
timeline and available activities, change contact information
and incentive preferences, complete study surveys, request a
self-collection kit for HIV testing, and report when they have
sent a specimen to the laboratory (see **Figure 2**). Participants
log into the app using a unique token that is generated at
enrollment. Once the app is downloaded, participants can set
their own PIN, which is required for subsequent entry to the
app. All collected data are encrypted in transit from the device
to the secure server, and identifying information is stored
encrypted within the database using the AES 256-bit protocol.
Staff with role-based access can access a dashboard to view
participant study progress, enter lab results, and view contact
information for shipment of specimen collection kits to
participants. All components of the app and other study materials
are available in the English and Spanish languages.

**Figure 2.** Participant view of the hybrid app of the online cohort, including (A) access to survey or HIV specimen collection kit within a study visit, (B) HIV test home kit shipping confirmation, and (C) home kit specimen collection instructions.

**Survey Measures**

All surveys are self-administered using the study hybrid app.
Survey measures span domains set forth by the social ecological
model [55] and by the situated vulnerabilities framework [56],
including individual, interpersonal, and structural level
measures. Unlike the site-based cohort, participants do not
complete a literacy screener and do not have an option for
staff-administered surveys, given the remote nature of data
collection.

Measures have been previously described [23] but briefly
include the following: individual-level measures collect data
on demographics; access to and use of gender-affirming services
[57]; sexual health history and access or uptake of HIV services
[58]; health insurance and primary care, mental and behavioral
health [59], and barriers and facilitators to care [60-62];
substance use [63-65]; gender affirmation and dysphoria [66];
and engagement in the HIVPC [67]. Interpersonal measures
include social support and social network size; detailed
partner-by-partner sexual behavior information by partner type
and HIV status [68,69]; and measures of intimate and nonpartner
physical, sexual, and psychological violence victimization [70].
Structural-level measures include social marginalization and stigma [71]; social capital [72,73]; interactions with the justice
system; and measures of immigration status, citizenship, and
interactions with immigration detention. As recommended by
the CAB, an additional measure of food insecurity was included
in the survey [74]. To mitigate risk of incomplete surveys due
to survey duration, longer scales that focused on gender pride,
community connectedness, and condom use self-efficacy are
not included in the online cohort survey.

At the end of the survey, information is provided to the
participant about nationally available and LGBTQ-friendly
locators for HIV testing, pre-exposure prophylaxis services, and
LGBTQ health centers. The survey is also programmed to
display locators for nationally available and LGBTQ-friendly
mental health services and suicide prevention; violence services;
and substance use services for individuals who report mental
health symptoms, experiences of violence victimization, or
substance use during the course of the survey. Both English and
Spanish language resources are included, though Spanish-speaking hotlines often have more limited hours.
Biologic Measures

Online cohort participants are asked to provide an oral fluid specimen at baseline and every 6 months for HIV testing. We elected to use oral fluid specimen collection, rather than home HIV self-testing, for 2 reasons: (1) preference to provide confirmatory testing, interactive posttest counseling, and support linkage to care for participants with positive or indeterminate preliminary results and (2) low self-reported response rates of HIV self-test results in other web-based HIV research [75,76]. Participants of the online cohort are not asked to participate in any STI testing or biospecimen storage, unlike site-based cohorts, due to resource constraints and added complexities of shipping specimens to multiple laboratories.

Specimen collection kits are sent by the study staff to the mailing address provided by the study participants. Participants are reminded to provide the name they would like shown on the package; they have the option to provide separate names for mailing vs email and email-based communication. Addresses are validated using the SmartyStreet application programming interface (API). Kits include the oral specimen collection device, OraSure, and easy-to-read bilingual information, including instructions for specimen collection with example photos, prelabeled shipping envelopes with tracking information, and a resource guide with locators for accessing HIV prevention or care resources and culturally appropriate health, mental health, substance use, and violence services.

Participants self-administer the OraSure Oral Specimen Collection Device at home in accordance with the package insert instructions [77]. The device allows for the collection of an oral fluid sample for HIV testing using a pad that is placed between the cheek and gum for 2-5 minutes. The pad is then placed into a sealable vial that contains a preservative, which stabilizes the sample for up to 21 days. Participants mail the vial and any corresponding materials to Quest Diagnostics using the prepaid label that is enclosed in the package. Participants are provided with a 24-hour OraSure Support telephone number as well as the study coordinator number if the participant has questions or difficulty using the collection device.

HIV testing is conducted at Quest Diagnostics Laboratory using an enzyme immune assay, with positive results being confirmed via western blot [78]. Confirmatory testing can take up to 6 additional days to complete. Results of the HIV tests are received by the study team within an average of 7 days. HIV test results are reviewed and provided to the participant within 48-72 hours of receipt via phone for positive results or via their preferred method of contact for negative results. The study staff review the data on a daily basis and call any participant with a positive or indeterminate test to provide posttest counseling and information about how to locate a provider in their location. All participants are informed during the consent process of the requirement to inform local health departments of positive HIV cases, which is performed by Quest Diagnostics per state requirement to inform local health departments of positive HIV test results.

Participants of the online cohort will contribute to the target sample size of 1100 transgender women who are not living with HIV in the LITE study. This target was based on a planned comparison of the incidence of 2 subgroups, assuming a two-sided test statistic and 5% Type 1 error with 80% power. We conservatively estimated 700 person-years of follow-up accumulated after the first 2 years of enrollment and, incorporating a loss of 15% of follow-up time per year, we estimated we will accumulate 935 person-years, 795 person-years, and 676 person-years in years 3, 4, and 5, respectively (>3000 person-years total). With this person-time, we estimated that we would have sufficient power (80%) power to detect public health relevant differences in HIV incidence by various participant characteristics. Relative to other online cohorts [37,39] and as one of the first cohorts among transgender women in the United States, we intentionally kept the sample size small to focus on developing acceptable methods and establishing trust among study participants.

Analytic Plan

Descriptive analysis will be used to compare differences in basic demographic, behavioral, and health-related characteristics among participants enrolled in the online vs the site-based cohorts, controlling for geographic region of residence.
Qualitative differences in recruitment and retention rates will also be observed and described.

Participant data from the online cohort will also contribute to the analysis of overall LITE study, which has been previously described [23]. Briefly, these analyses will include the following: Data visualization techniques will be utilized to compare efficiency of recruitment methods in terms of process measures of time, cost, and response rates, and descriptive analyses will be conducted to compare quantitative measures of acceptability, diversity, and biases of the samples recruited via technology-enhanced and traditional recruitment methods (Aim 1). Descriptive statistics will be conducted to compare demographic and behavioral risk profiles among baseline participants, as well as to assess baseline HIVCC and HIVPC among participants living with and without HIV, respectively (Aim 2). HIV incidence (and 95% confidence intervals) during follow-up will be estimated among the HIV-uninfected cohort (Aim 3). Poisson regression models will be used to investigate important predictors of incident infection. Time-to-event survival models will be used to estimate the risk of HIV seroconversion for the predictors. Finally, the steps in the HIVPC and HIVCC will be investigated over the course of the study using a serial cross-sectional approach to determine trends in these important frameworks (Aim 4). To make these estimates most informative for public health efforts, the steps will be estimated among those under observation in specific calendar years (ie, in 2019, 2020, and 2021).

Where relevant, we will utilize modern missing data techniques [79], with emphasis on multiple-imputation methods, which provide an approach for handling different missing data including coarsely measured [80] or mismeasured [81] variables. Adaptation of these methods to survival analyses will be similar to techniques implemented in R/MICE and SAS/IVEware [82-84]. Statistical analyses will be performed using Stata 14 (StataCorp, College Station TX) and R software (R Foundation for Statistical Computing, Vienna, Austria).

**Human Subjects**

Study activities follow a single institutional review board procedure and have undergone review and approval by the Johns Hopkins School of Medicine Institutional Review Board, with reliance by all collaborating organizations and institutions. Protection of participant privacy and confidentiality are central to the development and implementation of all study activities and have been further developed in discussions with the study CAB. The electronic data system has been developed with careful attention to security, including password protection and encrypted data transmission.

**Results**

In January 2019, we launched the online cohort in 25 cities, then expanded in April 2019 to a total of 72 cities in eastern and southern United States. In May 2019, we ended online advertisements, though eligible individuals who had heard of the study from other participants could continue to enroll. Enrollment closed in August 2020, at which point 580 participants were fully enrolled in the online cohort. We observed a recruitment-enrollment cascade in the study implementation process with successive and decreasing rates from online screening through enrollment in the online cohort: 2387 individuals were screened, 1092 individuals consented, 971 individuals completed the survey, 586 individuals completed both the HIV test and survey, and 580 individuals enrolled in the cohort (Figure 3). Ultimately, 45.7% (1092/2387) of those screened were consented, and 53.3% (580/1092) of those consented completed both baseline study activities for full enrollment in the cohort. The self-administered screening form effectively identified ineligible participants on the basis of HIV status, minimizing HIV testing costs, where only 4 participants were newly identified with HIV infection through baseline HIV testing procedures. Discussions with these participants determined that 3 of the 4 were newly diagnosed with HIV, and staff were able to provide appropriate counseling and referral to HIV care for these participants.

**Figure 3.** Recruitment-enrollment cascade: progression from screening to enrollment of an online cohort of transgender women.
Figure 4 displays cumulative enrollment over time into the online cohort. Over one-third (213/580, 36.8%) of the online sample were aged 18-24 years at enrollment, and one-fifth (125/580, 21.6%) identify as Black, Latinx, or multiracial. Across site-based and online cohorts, we have enrolled and follow 1312 transgender women at risk for HIV in eastern and southern United States. As of the end of October 2020, 573 person-years have been contributed to the analytic dataset of the online cohort. Cohort participants are under follow-up through 2022.

Figure 4. Cumulative number of enrolled participants in the LITE online cohort.

Internal shipping procedures for specimen collection kits have also led to observations of high mobility or unstable housing among cohort participants when specimen collection kits are returned by mail. While requiring more staff effort, returned packages offer an opportunity to identify address changes and connect with participants to support retention. Returned specimen collection kits during enrollment also supported identification of some fraudulent enrollments.

Discussion

This paper describes the protocol for the development and implementation of an online cohort for assessment of HIV risks and predictors of HIV acquisition among transgender women. In conjunction with the LITE site-based cohort, this study aims to inform priorities and programs of the Ending the HIV Epidemic Strategy for the United States [1,2] and innovates in its use of technology to engage and assess HIV risk among transgender women. Study findings will have important implications for identification of optimal recruitment, retention, and data collection methods for future observational research and digital or hybridized digital clinical trials among transgender women living with and without HIV infection, estimation of HIV incidence, and ultimately the development of acceptable HIV prevention strategies for ending the HIV epidemic among transgender women in the United States. As with other online cohorts and cross-sectional studies [37,39,75], we noted a recruitment-enrollment cascade from screening through enrollment and completion of study activities, which has important design, budgetary, and logistical implications for future digital epidemiologic studies and clinical trials.

This online cohort is strengthened by extensive formative research and early experiences incorporating technology-enhanced methods into the site-based cohorts [22,23,47,48]. Though we make every effort to ensure that technology-enhanced and digital methods are accessible, there is a risk that individuals with extremely limited access to technology may be excluded from this online cohort. Thus, the rare ability to compare study findings across site-based and online cohorts provides invaluable information about sources of bias, but more importantly indicates if use of digital methods risks excluding subgroups who could most benefit from HIV research and subsequent HIV prevention and care interventions.
Acknowledgments
The authors would like to express their gratitude to the transgender women who take part in this study. This study would not be possible without their participation. We appreciate the continued involvement and contributions by the CAB that supports and guides this study. Research reported in this publication was jointly supported by the National Institute of Allergy and Infectious Diseases, National Institute of Mental Health, and National Institute of Child Health and Human Development of the National Institutes of Health under Award Number UG3/UG1AI133669 (Wirtz/Reisner). Research reported in this publication was also supported by HIV/AIDS, Hepatitis, STD, and TB Administration (HAHSTA), Washington DC Department of Health. The LITE study is also appreciative of support from the CFAR at partner institutions, including JHU (P30AI094189), Emory University (P30AI050409), Harvard University (P30AI060354), DC CFAR (AI117970), and the University of Miami (P30AI073961). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or HAHSTA.

Authors’ Contributions
The following are members of the collaborative author, American Cohort to Study HIV Acquisition Among Transgender Women (also known as the LITE cohort): Sari Reisner (multiple principal investigator [PI]; Harvard University, Brigham and Women's Hospital); Andrea Wirtz (multiple PI; Johns Hopkins University [JHU]); Keri Althoff (JHU); Chris Beyrer (JHU); James Case (JHU); Erin Cooney (JHU); Oliver Laeyendecker (JHU); Tonia Poteat (University of North Carolina); Ken Mayer (Fenway Health); Asa Radix (Callen-Lorde Community Health Center); Christopher M Cannon (Whitman-Walker Institute); Jason Schneider (Emory University and Grady Hospital); J. Sonya Haw (Emory University and Grady Hospital); Allan Rodriguez (University of Miami); Andrew Wawrzeniak (University of Miami); and the LITE CAB, including Sherri Meeks, Sydney Shackelford, Nala Toussaint, SaVanna Wanzer, as well as those who have remained anonymous. ALW and SLR developed the study concept; TP, AR, KNA, CC, AW, EC, and KM provided extensive input to the original grant submission and/or the protocol manuscript; ALW wrote the first draft of the manuscript; MS coordinates the online cohort; and all authors reviewed and provided scientific input to the manuscript.

Conflicts of Interest
KNA serves as a consultant to the All of Study (NIH) and on the scientific advisory board for TrioHealth Inc.

References


25. Black Women’s Health Study. Boston University Slone Epidemiology Center. URL: https://www.bu.edu/bwhs/history/ [accessed 2021-04-05]


https://www.researchprotocols.org/2021/4/e29152


Abbreviations

API: application programming interface
CAB: community advisory board
HIVCC: HIV Care Continuum
HIVPC: HIV Prevention Continuum
JHU: Johns Hopkins University
PI: principal investigator
STI: sexually transmitted infection

©Andrea L Wirtz, Erin E Cooney, Megan Stevenson, Asa Radix, Tonia Poteat, Andrew J Wawrzyniak, Christopher M Cannon, Jason S Schneider, J Sonya Haw, James Case, Keri N Althoff, Elizabeth Humes, Kenneth H Mayer, Chris Beyrer, Allan E Rodriguez, Sari L Reisner, American Cohort To Study HIV Acquisition Among Transgender Women (LITE) Study Group

Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.04.2021. This is an open-access article
distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.
The Geriatric Acute and Post-Acute Fall Prevention Intervention (GAPcare) II to Assess the Use of the Apple Watch in Older Emergency Department Patients With Falls: Protocol for a Mixed Methods Study

Daniel H Strauss1, BS; Natalie M Davoodi2, BS; Margaret Healy3, BS; Christopher L Metts4, MD; Roland C Merchant5, MD, MPH, ScD; Swechya Banskota1, BS; Elizabeth M Goldberg6, MD, ScM

1Alpert Medical School of Brown University, Providence, RI, United States
2Brown University School of Public Health, Providence, RI, United States
3Rhode Island Hospital, Providence, RI, United States
4Department of Pathology and Laboratory Medicine, Medical University of South Carolina, Charleston, SC, United States
5Department of Emergency Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, United States
6Department of Emergency Medicine, Alpert Medical School of Brown University, Providence, RI, United States

Corresponding Author:
Elizabeth M Goldberg, MD, ScM
Department of Emergency Medicine
Alpert Medical School of Brown University
55 Claverick Street, Second Floor, Room 203
Providence, RI, 02903
United States
Phone: 1 401 527 1740
Email: elizabeth_goldberg@brown.edu

Abstract

Background: Falls are a common problem among older adults that lead to injury, emergency department (ED) visits, and institutionalization. The Apple Watch can detect falls and alert caregivers and clinicians that help is needed; the device could also be used to objectively collect data on gait, fitness, and falls as part of clinical trials. However, little is known about the ease of use of this technology among older adult ED patients, a population at high risk of recurrent falls.

Objective: The goal of this study—the Geriatric Acute and Post-Acute Fall Prevention Intervention (GAPcare) II—is to examine the feasibility, acceptability, and usability of the Apple Watch Series 4 paired with the iPhone and our research app Rhode Island FitTest (RIFitTest) among older adult ED patients seeking care for falls.

Methods: We will conduct field-testing with older adult ED patients (n=25) who sustained a fall and their caregivers (n=5) to determine whether they can use the Apple Watch, iPhone, and app either (1) continuously or (2) periodically, with or without telephone assistance from the research staff, to assess gait, fitness, and/or falls over time. During the initial encounter, participants will receive training in the Apple Watch, iPhone, and our research app. They will receive an illustrated training manual and a number to call if they have questions about the research protocol or device usage. Participants will complete surveys and cognitive and motor assessments on the app during the study period. At the conclusion of the study, we will solicit participant feedback through semistructured interviews. Qualitative data will be summarized using framework matrix analyses. Sensor and survey response data will be analyzed using descriptive statistics.

Results: Recruitment began in December 2019 and was on pause from April 2020 until September 2020 due to the COVID-19 pandemic. Study recruitment will continue until 30 participants are enrolled. This study has been approved by the Rhode Island Hospital Institutional Review Board (approval 1400781-16).

Conclusions: GAPcare II will provide insights into the feasibility, acceptability, and usability of the Apple Watch, iPhone, and the RIFitTest app in the population most likely to benefit from the technology: older adults at high risk of recurrent falls. In the future, wearables could be used as part of fall prevention interventions to prevent injury before it occurs.

Trial Registration: ClinicalTrials.gov NCT04304495; https://clinicaltrials.gov/ct2/show/NCT04304495
**Introduction**

Falls are the leading cause of emergency department (ED) visits for injuries in older adults aged 65 and older [1,2]. Each year in the United States, 28.7% of older adults sustain a fall, resulting in 2.8 million annual ED visits for geriatric falls [3]. The risk of recurrent falls is particularly high immediately after the ED visit [4]; 31% of community-dwelling older adults in the United States fall again within 6 months, and 62% of these falls cause serious injury [5,6]. Technology-based interventions initiated immediately after the ED visit could be useful to prevent falls, if they are embraced by older adults.

Wearable technology, while conventionally targeted toward a younger population, has shown promise in older adults. One recent prospective study with 95 community-dwelling older adults found as many as 91% of participants rated a wearable watch as acceptable and easy to use [7]. Additionally, the Apple Heart Study, launched in 2017, recruited over 24,000 older adults to monitor their heart rhythm for atrial fibrillation using the Apple Watch and an iPhone app [8], indicating that this population can use these devices. Furthermore, a recent systematic review showed that exercise, compared to educational or environmental interventions, significantly reduced falls in older adults, although increased walking alone did not decrease falls [9]. Wrist-worn technologies may increase daily step counts [10] and encourage exercise tracking, which could facilitate fall prevention. Fitness tracking has the additional psychological benefit of allowing individuals to have control of their subjective well-being, how people experience and evaluate the quality of their lives [11]. The prospect of these devices for fall-related measures in older adults remains a relatively untapped area for application of this technology.

The Apple Watch (Apple Inc) is a wearable device with built-in sensors that can detect falls automatically using its accelerometer and gyroscope sensors. Unlike traditional fall alert pendants, the Apple Watch alerts designated contacts and emergency medical services even if the individual is unconscious or immobile. The Apple Watch is particularly well-suited as a research tool because fall occurrences and other sensor data (eg, heart rate and step count) are recorded in HealthKit, a developer application programming interface included in the iPhone operating system (iOS) [12]. Additionally, Apple provides developers with open-source code for cognitive and motor tests called *active tasks*, which allow participants to perform assessments independently at home with the help of visual, tactile, and auditory prompts generated by the iPhone. Results of these assessments, as well as fall occurrences and other HealthKit-collected data, are recorded into our novel research app Rhode Island FitTest (RIFitTest) and transmitted to Research Electronic Data Capture (REDCap) [13] in real time for monitoring and analysis by the research team.

We used the Unified Theory of Acceptance and Use of Technology (UTAUT) model as a conceptual framework when developing our protocol. The UTAUT model, which is a combination of eight prominent theories—the Theory of Reasoned Action, Innovation Diffusion Theory, the Theory of Planned Behavior (TPB), the Technology Acceptance Model (TAM), the combined TAM-TPB, Social Cognitive Theory, the Motivational Model, and the Model of Personal Computer Utilization—has been extensively tested across multiple disciplines and has been validated as an effective tool to assess acceptance of health-related technology [14]. In the UTAUT model, four constructs play a significant role as determinants of user acceptance and usage of technology: (1) performance expectancy, (2) effort expectancy, (3) self-efficacy, and (4) facilitating conditions.

In this mixed methods study—the Geriatric Acute and Post-Acute Fall Prevention Intervention (GAPcare) II—we will conduct field-testing with older adult ED patients, with and without cognitive impairment, who present to the ED with a fall within the last 7 days, as well as with their caregivers, to assess the feasibility, acceptability, and usability of the Apple Watch, iPhone, and the RIFitTest app to assess gait, fitness, and/or subsequent falls.

**Methods**

**Study Design**

Because we are interested in reducing recurrent falls among older adult ED patients, we designed a mixed methods study to collect perspectives from older adults and their caregivers through semistructured interviews (ie, qualitative data) and information about falls and fall risk factors prospectively (ie, quantitative data). Fall risk factors include impairment in cognition, impairment in mobility, fear of falling, and heart rate abnormalities, all of which can be measured by wearable device sensors or through survey questions on mobile devices. This study was registered at ClinicalTrials.gov (NCT04304495).

**Field-Testing Nature of the Study**

Older adults face unique challenges with using mobile devices, and many studies have shown that interest in using wearables often wanes after initial recruitment [15]. Therefore, we decided we needed to first field-test our protocol among older adults to ensure they would express interest in participating, be able to learn how to use the technology, and stay continuously engaged in the study. A field test is a type of testing that evaluates technology performance in real-world settings. This allows for an enhanced understanding of how the target end user experiences the product and can lead to design optimization and insights into challenges that must still be overcome.
Planning Phase

During the planning phase, we assembled experts in digital health, qualitative research, geriatrics, neurocognitive assessments, and clinical research. This team advised the principal investigator (EG) on the protocol, outcome assessment, and implementation of the study. We also programmed our survey questions in REDCap, created the app, and tested the app among our research staff during this time.

Information Technology Solution

To collect the sensor and survey data generated on the Apple Watch and to transfer the data to our secure server, we needed to develop an iOS app. We engaged one of the study authors (CM) to assist with creating the app for the study using his third-party platform, status/post [16]. Status/post integrates REDCap and ResearchKit and allows researchers to design the app in the same way questionnaires are designed in REDCap, reducing costs and turnaround time because a developer is not necessary.

RIFitTest Study App

This app, when paired with the Apple Watch and the user’s iPhone, collects sensor-obtained data passively. Users can also enter data actively using their phones. Several steps are required to create the research app, RIFitTest:

1. Enter survey questions and answers into REDCap data collection software.
2. Choose Apple HealthKit measures (eg, heart rate and steps) and enter Apple Inc open-source code into the field notes of REDCap to extract sensor data from the watch and store it in REDCap.
3. Select Apple ResearchKit active tasks relevant to the study and enter Apple Inc open-source code into the field notes of REDCap.
4. Use status/post to schedule notifications for participants, create participant usernames and passwords, and monitor participant activity.
5. Post app on Apple App Store to allow participants to download the app onto their iPhones.

After creation of the app, we will pilot-test it among the research staff and make changes to improve the usability for older adults (eg, increase font size and choose the color of the app to enhance readability). Only participants who complete the consent process will be able to download the study app. When installing the RIFitTest app, participants will enter a username and password created by the study team and will permit the app to access each category of data stored in Apple HealthKit (eg, heart rate and step count).

Data Security, Transmission, and Storage

Data from the device will be pushed to the Health Insurance Portability and Accountability Act (HIPAA)-compliant REDCap program when participants connect to a Wi-Fi network, or the data will be added to the REDCap program manually during follow-up with the research team. We will use an application programming interface to ensure data can be viewed in real time on the REDCap program. Data transmission is encrypted with Secure Sockets Layer. Similar to other iOS-based apps, the RIFitTest app is “sandboxed,” meaning that no other apps can gather data collected by the study app. The app is password protected, which prevents user data from being revealed if the device is stolen.

Study Protocol

Setting

The study will be conducted at two EDs in Providence, Rhode Island: Rhode Island Hospital (RIH), a tertiary-care hospital, and The Miriam Hospital (TMH), an academic community hospital. RIH is the only Level I trauma center in the state, while patients at TMH are primarily community-dwelling older adults.

Population

We will recruit 30 participants. These participants will be recruited into six groups with 5 participants each from two different cohorts: cognitively intact and cognitively impaired. Cognitively intact participants will include 20 patients: ages 65-69 (n=5), ages 70-74 (n=5), ages 75-79 (n=5), and ages 80-84 (n=5). Cognitively impaired participants aged 65 years old and older (n=5) will be enrolled with their caregivers (n=5). Caregivers will complete the same tasks as the patient and are encouraged to assist the patient in using the technology. Their perspectives are important for understanding how the technology can be used for communication and to better understand facilitators and barriers of the technology. Quota sampling will be employed to ensure study participants reflect the racial and ethnic diversity of Rhode Island.

Eligibility

Eligible participants will be noninstitutionalized, community-dwelling older adults aged 65 years and older who are English-speaking and present to the ED after a fall within the past 7 days. Patients with cognitive impairment, as measured by a score of less than 4 on the Six-Item Screener [17], must have a legally authorized representative available to give informed consent. The patient’s ED physician must intend to discharge the patient after their initial evaluation either to their home or to an assisted living or rehabilitation center. Patients with falls due to syncope, an externally applied force, or critical illness (eg, stroke) will be excluded. Patients who present with altered mental status (eg, intoxicated or agitated), have injuries that prevent mobilization (eg, pelvic or lower extremity fractures), are undomiciled, have allergies to any wearable device component, are unable or unwilling to wear an Apple Watch at home, or have advanced cancer and/or are in hospice care also will be excluded.

Recruitment and Enrollment

Eligible patients will be approached by a member of the research team and asked if they are interested in participating. If interested, participants will provide written informed consent. Those who consent will answer questions on prior technology use, demographic characteristics, previous falls, and their health history. Participants will receive an Apple Watch Series 4 and an iPhone 7—if they do not already own these devices—for the duration of the study. Participants will be able to use their own Apple Watch and iPhone for the study.
Apple Watch Specifications and Deployment
For this study, each participant will use the Apple Watch Series 4 (GPS), manufactured by Quanta Computer Compal Electronics. We will use the Apple Watch with the larger 44-mm watch face and Velcro wrist strap to make it easier to use for older adults with vision impairment or those lacking fine motor skills.

ED Procedures
The following steps will be taken for device setup prior to patient approach:
1. Charge and pair the Apple Watch and iPhone via Bluetooth; for participants with their own devices, pairing will occur in real time.
2. Maximize font and icon size on both devices to enhance readability.
3. Set brightness at its maximum to improve readability.
4. Delete irrelevant content and minimize notifications (eg, remove and mute unnecessary apps and notifications, respectively) to limit distractions.
5. Download the RIFitTest app from the Apple App Store onto the iPhone.
6. Add the study principal investigator and an emergency contact to the contact list to receive fall alert notifications.

Training in the use of the iPhone, Apple Watch, and RIFitTest app as well as study-related activities will also be provided, including the following:
1. iPhone and Apple Watch setup instructions, including how to turn the device on and off, unlock and set a password, and charge both devices.
2. Real-time demonstration of active tasks accessed through the study app.
3. A demonstration of how to log a fall and/or injury through the app.
4. Education in the research protocol, including when to use the Apple Watch (ie, continuously, including while bathing and participating in water activities), nightly charging of iPhone, mealtime charging of the Apple Watch, daily completion of fall surveys, and weekly completion of active tasks on the RIFitTest app.
5. Orientation to the study materials: file, business card with hotline for assistance, technology manual (see Multimedia Appendix 1), calendar of assessment timing, technology agreement, and consent form.

Figure 1 illustrates a timeline of the study that will be explained to participants.
Follow-up Procedures
Study participation will last for 30 days or until the participant is no longer able or willing to participate. All participants will be contacted by research staff on days 3, 8, 15, and 22 to remind them to complete the steps in the study protocol and to answer questions. Participants will take part in a 30-minute semistructured interview at the completion of the study period. The interview will be conducted by the principal investigator and her staff. The intent of the interview will be to record participants’ experience with the study as well as to provide information on the feasibility, usability, and acceptability of the Apple Watch, iPhone, and RIFitTest app. Participant feedback about the quality of the intervention materials will also be solicited. The information collected in the interview will be used to improve the app and training procedures in preparation of the subsequent clinical trial.

Measures
There are five data sources that will contribute data for this field test: research staff–administered surveys, qualitative interviews with patients and caregivers, Apple Watch sensor-obtained measures, user-completed iPhone surveys (ie, a digital fall diary, fall efficacy survey, and a usability survey), and user-completed active tasks (ie, guided motor and cognitive assessments). Specific measures are summarized in Table 1.
Table 1. Study instruments, surveys, and active tasks.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Description</th>
<th>Administration time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Process evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research staff–administered survey</td>
<td>This survey will record how many patients were screened, agreed to participate, were recruited, received intended treatment, and were retained.</td>
<td>Baseline and 30 days</td>
</tr>
<tr>
<td><strong>Outcome evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research staff–administered surveys and interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment questionnaire</td>
<td>Demographic characteristics</td>
<td>Baseline; &lt;5 minutes</td>
</tr>
<tr>
<td></td>
<td>Prior fall history, comorbidities, emergency department index visit fall circumstances, and injuries</td>
<td></td>
</tr>
<tr>
<td>Six-Item Screener [17]</td>
<td>6-point questionnaire to measure cognitive impairment for study screening; &lt;4 indicates high risk for cognitive impairment</td>
<td>Baseline; 2 minutes</td>
</tr>
<tr>
<td>Qualitative interviews</td>
<td>These interviews will track follow-up phone sessions with research staff. Interviews record participants’ experience with the study while also establishing an understanding of the usability of both the Apple Watch and RIFitTest&lt;sup&gt;a&lt;/sup&gt; app for older adults. The timing and length of each session and barriers to attendance are recorded.</td>
<td>Days 3, 8, 15, 22, and 30</td>
</tr>
<tr>
<td><strong>Apple Watch sensor-obtained measures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple Watch measures (ie, accelerometer, gyroscope, and physiological sensors)</td>
<td>Gait and fitness (ie, time spent walking, standing, and climbing steps; 4-meter walk gait speed test; and resting heart rate) Data will include cognitive and active task performance, including cadence and walking speed related to their gait and balance and Stroop test. Performance reports will be generated for each participant.</td>
<td>Baseline ± continuously (days 1 to 30)</td>
</tr>
<tr>
<td><strong>User-completed iPhone surveys</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall diary</td>
<td>Daily app-based entry that will ask participants if they have fallen</td>
<td>Days 1 to 30</td>
</tr>
<tr>
<td>Fall-efficacy survey</td>
<td>Measure of fear of falling, which increases fall risk</td>
<td>Days 1 and 28</td>
</tr>
<tr>
<td>Usability survey</td>
<td>This survey will record each participant’s experience with the app, as well as with the active and cognitive tasks.</td>
<td>Days 7 and 28</td>
</tr>
<tr>
<td><strong>User-completed active tasks [12] on the iPhone</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gait and balance test</td>
<td>Active task; measure of 20 steps of walking in one direction, returning, and then standing still for 30 seconds</td>
<td>Days 1, 7, 14, 21, and 28</td>
</tr>
<tr>
<td>Timed walk</td>
<td>Measure of a walking distance of 109 yards in a straight line as quickly as possible</td>
<td>Days 1, 7, 14, 21, and 28</td>
</tr>
<tr>
<td>Stroop test</td>
<td>Measure of ability to quickly identify color of text</td>
<td>Days 1, 7, 14, 21, and 28</td>
</tr>
<tr>
<td>Reaction time test</td>
<td>Active task; measure of speed at clicking a button when shown</td>
<td>Days 1, 7, 14, 21, and 28</td>
</tr>
<tr>
<td>Trail-making test</td>
<td>Active task; measure of ability to follow an alphanumeric sequence (1...A...2...B..., etc)</td>
<td>Days 1, 7, 14, 21, and 28</td>
</tr>
</tbody>
</table>

<sup>a</sup>RIFitTest: Rhode Island FitTest.

**Qualitative Interviews**

**Interview Procedure**

Study staff will conduct interviews with both the participants and their caregivers, when applicable. They will use an interview guide containing semistructured questions with follow-up questions and probes to explore participants’ responses.

**Recording**

Interviews will be transcribed and deidentified. Transcripts will be corrected for accuracy as needed using the audio recording.

**Analysis**

We will use framework analysis—a qualitative analysis technique, in which investigators summarize content within categories into charts after transcription [18,19]—as it is particularly well-suited to generate recommendations within a
limited time period, so as to inform the subsequent clinical trial. Another benefit of the framework method is the emphasis on transparency in data analysis and the links between the stages of the analysis [19]. Framework methods include summarizing qualitative data into a matrix or spreadsheet where data are entered by codes (columns) and cases (rows) [20]. Specifically, we will take the following steps:

1. Read and re-read the transcripts to note initial observations; iteratively search for common themes, subthemes, and patterns across participant responses; and develop a coding matrix and assign data to the themes and categories in the coding matrix.
2. Review themes in relation to the extracts and entire data set; identify associations between the themes until a “whole picture” is apparent.
3. Select representative quotes from the interviews to illustrate the themes.
4. Record ideas about emerging themes in an ongoing audit trail.
5. Prepare the analytic narrative and contextualize it using existing research on this topic.

Debriefing

Interviewers will complete a debriefing form after the interview to document the following: (1) tone of interview, (2) agenda adherence, (3) interview description, (4) major themes, (5) lessons learned, (6) question strategies, and (7) saturation.

Poststage Interview Content

The following three study aims will be addressed through qualitative data collected during the poststage interview:

1. Feasibility: (1) ability to carry out device charging, application, and manipulation; (2) barriers to daily use; and (3) responses to specific planned components of the intervention.
2. Acceptability: receptivity to and concerns about device use.
3. Usability: (1) prior experiences with wearables; (2) trajectories of use, including reasons for continuation or cessation of use; (3) quality of in-person and telephone device support and training; (4) connectivity to smartphone and Wi-Fi, if desired; (5) caregiver experiences; and (6) patient and caregiver’s perceived needs and preferred method of use.

Interview questions will be grounded in the four constructs of the UTAUT model.

Sensor and Survey Data Collection and Analysis

Continuous data collected by the Apple Watch will be sourced from the accelerometer, gyroscope, heart rate monitor, and GPS sensors. Episodic data that will be collected include electronic health record data, fall dairy, Apple Watch measures, RIFitTest sensors. Episodic data that will be collected include electronic health record data, fall dairy, Apple Watch measures, RIFitTest sensors. Episodic data that will be collected include electronic health record data, fall dairy, Apple Watch measures, RIFitTest sensors.

Descriptive statistics will be used to measure the following key parameters of feasibility of recruitment, enrollment, participation, and retention [21]: number of patients screened, eligible, and recruited; time required to recruit; number of patients unable to provide consent; number of dropouts; and number of patients with refusal and retention at each follow-up. Frequencies, proportions, rates, means and medians, standard deviations, and other measures of variability will be used to report on these feasibility measures.

Descriptive statistics will also be used to assess the following parameters of acceptability: the fall diary completion rates, days the Apple Watch was worn, and any malfunctions of the Apple Watch or app. We will use the Apple Watch–obtained data to determine the association between gait and fitness measures (eg, step count, stairs climbed, and resting heart rate) and reported falls. To account for the longitudinal nature of this data and within-person correlation, we will use generalized estimating equations or generalized linear models [22].

Privacy and Data Storage

The Apple Watch and iPhone will be password protected. All devices will be cleaned following hospital procedures and the stored data will be erased upon return of these devices at the end of the participants’ involvement in the study.

Results

Recruitment began in December 2019 and was on pause from April 2020 until September 2020 due to the COVID-19 pandemic. Funding for this work includes a K76 grant awarded by the National Institute on Aging (NIA), via the Paul B Beeson Emerging Leaders Career Development Award in Aging (grant K76AG059983), for the period of September 1, 2019, to September 1, 2024. This study has been approved by the RIH Institutional Review Board (approval 1400781-16).

Discussion

Overview

GAPcare II will field-test the Apple Watch, iPhone, and our novel research app RIFitTest among older adults. Our study may open new research horizons by providing data on how to implement sensor-based assessment in real-world settings in this population. The successful refinement and implementation of the protocol could provide a framework to help other scientists leverage sensor technologies in the older adult population. A growing number of clinical trials will incorporate sensor-based wearables in the next decade [23], and our protocol describes how to field-test these devices among the patient population before initiating a clinical trial.

Technology-based interventions hold great promise in the field of fall prevention. Fall detection and medical alert devices can collect and transmit health information about older adults in a timely fashion directly to investigators or clinicians. This data could be used to help researchers and clinicians prevent future falls. However, current studies assessing the feasibility, acceptability, and usability of the Apple Watch in the population must at risk for falls are lacking. Although the Apple Watch is already available to consumers for recreational use, modifications to the settings and app creation are necessary to collect data securely from users and ensure optimal use by older
individuals. The GAPcare II protocol may be useful to help researchers who are planning to recruit patients during the COVID-19 pandemic who may be confined to their homes. Our protocol describes how participants can be recruited and can contribute valuable information about cognition and motor skills by using an app from home. While there are a growing number of clinical trials using sensor-based technology, best practices for deploying these technologies among older adults have not yet been established. This field test may help guide the successful integration of wearable devices into clinical studies in this population.

The Apple Watch’s multifunctionality is useful because it combines acceleration-based fall detection methods with caregiver and medical personnel communication. Traditional fall monitoring technology often requires manual activation of a button, which is impossible for a person who may be unable to get up independently after a fall or could be unconscious, and many older adults feel a stigma is associated with wearing a device that is solely for fall alerts. The Apple Watch and other fitness trackers with functions other than fall alerts may have greater appeal for older adults who are seeking to improve their health and stay in communication with loved ones.

While digital health technologies such as wearables show promise, they are expensive and may not be accessible to all older adults. One study conducted in a nationally representative sample of adults in the United States found that patients with low health literacy are less likely to use digital health tools. Conversely, the same study showed that adults with adequate health literacy are more likely to use wearable technologies such as activity trackers, and are more likely to report finding them useful and easy to use [24]. These findings suggest a need to improve usability and functionality of these technologies for all older adults, especially those with low health literacy. Moreover, racial and ethnic minorities use health-related technology less than White older adults, and older adults with higher incomes and greater educational attainment use health-related technology more frequently [25-27]. However, some Medicare Advantage companies are now subsidizing the cost of devices to encourage fitness, which could make these wearables more accessible [28]. To ensure equitable access to wearable technology, other payors should consider subsidizing the cost of these devices that may otherwise be financially inaccessible to many people nationwide. Wearable technology remains inaccessible to many older adults, and efforts to broaden technology literacy and increase cost-effectiveness is of vital concern.

For all older adults, internet and wearable technologies may lack appeal due to concerns about privacy and safety [29], limiting the potential scope and use of Apple Watch technology. Furthermore, there is resistance to uptake and long-term use of wearable technology for fall detection and medical alert in the older adult population, with studies showing that even if a user had an overall positive first impression of the device, it did not lead to long-term device use [30]. The Apple Watch, as well as other Apple products, may appeal more to a younger audience, which may reduce uptake and long-term viability in older populations. For this reason, users need to be educated on the benefits of the maintenance of device use and must receive training to build their internal motivation for device use. Our training protocol and manual could be useful for this purpose. Although drawbacks exist to wearable technology, home-based care is of rising importance, and this technology may improve self-sufficiency among older adults, while also providing important research data to help prevent falls.

Limitations

While GAPcare II may provide important insights into older adults and their technology preferences, there are several limitations. The sample size of 30 participants is not powered or designed to detect cognitive and motor functioning trajectories and the accuracy of the Apple Watch fall algorithm, but rather to field-test the use of these devices in older adults. Because this study will be conducted in an urban setting among English-speaking ED patients, study results may not be generalizable to rural populations or adults outside of New England. Bilingual research staff are unavailable to conduct recruitment, follow-up, and final interviews; however, expansion among non-English speakers will be paramount in future work to ensure a diverse participant pool. Apple products, including the Apple Watch and iPhone 7, may be cost-prohibitive for patients that represent marginalized populations. We provide iPhones on loan in our study to overcome this barrier. The Apple Watch does not integrate with non-Apple products, which may be an additional challenge individuals need to overcome as they learn how to use these devices for health purposes.

Conclusions

GAPcare II will provide insights into the feasibility, acceptability, and usability of the Apple Watch, iPhone, and the RIFitTest app in older adults who seek care for falls. Mobile technology could be used as part of clinical trials to objectively measure fall outcomes and could provide detailed continuous information on cognitive and motor functioning without the need for guided assessments by in-person research staff. In the future, wearables could be used as a part of fall prevention interventions to prevent injury before it occurs.

Acknowledgments

We would like to acknowledge the Emergency Medicine Clinical Research teams at TMH and RIH for their assistance in recruitment and data collection for this study, DS, ND, and MH contributed to writing the manuscript, editing, and redrafting. SB wrote portions of the paper and provided critical editing. EG devised, wrote, and organized the contributions of this manuscript. CM and RM provided critical edits and revisions in writing the manuscript. EG received funding for this work from the NIA (grant K76AG059983). Apple Watches were supplied for study purposes from Apple through Apple’s Investigator Support

https://www.researchprotocols.org/2021/4/e24455
Program. SB received funding for this work from the NIA, via the Medical Student Training in Aging Research (MSTAR) Program (grant 5T35AG038027-10), and the American Federation for Aging Research.

Conflicts of Interest
CM is the developer of status/post through Infinite Arms, LLC.

Multimedia Appendix 1
Instructional manual given to participants.

References
12. Active tasks. Apple ResearchKit. URL: http://researchkit.org/docs/docs/ActiveTasks/ActiveTasks.html [accessed 2020-09-20]


Abbreviations

ED: emergency department
GAPcare: Geriatric Acute and Post-Acute Fall Prevention Intervention
HIPAA: Health Insurance Portability and Accountability Act
iOS: iPhone operating system
MSTAR: Medical Student Training in Aging Research
NIA: National Institute on Aging
REDCap: Research Electronic Data Capture
RIFitTest: Rhode Island FitTest
RH: Rhode Island Hospital
TAM: Technology Acceptance Model
TMH: The Miriam Hospital
TPB: Theory of Planned Behavior
UTAUT: Unified Theory of Acceptance and Use of Technology

Please cite as:
Strauss DH, Davoodi NM, Healy M, Metts CL, Merchant RC, Banskota S, Goldberg EM
The Geriatric Acute and Post-Acute Fall Prevention Intervention (GAPcare) II to Assess the Use of the Apple Watch in Older Emergency Department Patients With Falls: Protocol for a Mixed Methods Study
JMIR Res Protoc 2021;10(4):e24455
URL: https://www.researchprotocols.org/2021/4/e24455
doi:10.2196/24455
PMID:33792553
Protocol

Development and Validation of Clinical Prediction Models for Surgical Success in Patients With Endometriosis: Protocol for a Mixed Methods Study

Nadine Marlin¹, DipSBM, MSc; Carol Rivas², BSc (Hons), MSc, PhD; John Allotey¹, BSc, MSc, PhD; Julie Dodds³, BSc, PhD; Andrew Horne⁴, MBChB, PhD; Elizabeth Ball³, MD, PhD

¹Pragmatic Clinical Trials Unit, Centre for Primary Care and Public Health, Queen Mary University of London, London, United Kingdom
²UCL Social Research Institute, University College London, London, United Kingdom
³Women's Health Research Unit, Centre for Primary Care and Public Health, Queen Mary University of London, London, United Kingdom
⁴MRC Centre for Reproductive Health, University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author: Carol Rivas, BSc (Hons), MSc, PhD
UCL Social Research Institute
University College London
18 Woburn Square
London, WC1H 0NR
United Kingdom
Phone: 44 2076126923
Email: c.rivas@ucl.ac.uk

Abstract

Background: Endometriosis is a chronic inflammatory condition affecting 6%-10% of women of reproductive age and is defined by the presence of endometrial-like tissue outside the uterus (lesions), commonly affecting the pelvis and ovaries. It is associated with debilitating pelvic pain, infertility, and fatigue and often has devastating effects on the quality of life (QoL). Although it is as common as back pain, it is poorly understood, and treatment and diagnosis are often delayed, leading to unnecessary suffering. Endometriosis has no cure. Surgery is one of several management options. Quantifying the probability of successful surgery is important for guiding clinical decisions and treatment strategies. Factors predicting success through pain reduction after endometriosis surgery have not yet been adequately identified.

Objective: This study aims to determine which women with confirmed endometriosis benefit from surgical improvement in pain and QoL and whether these women could be identified from clinical symptoms measured before laparoscopy.

Methods: First, we will carry out a systematic search and review and, if appropriate, meta-analysis of observational cohort and case-control studies reporting one or more risk factors for endometriosis and postsurgical treatment success. We will search PubMed, Embase, and Cochrane databases from inception without language restrictions and supplement the reference lists by manual searches. Second, we will develop separate clinical prediction models for women with confirmed and suspected diagnoses of endometriosis. A total of three suitable databases have been identified for development and external validation (the MEDAL [ISRCTN13028601] and LUNA [ISRCTN41196151] studies, and the BSGE database), and access has been guaranteed. The models will be developed using a linear regression approach that links candidate factors to outcomes. Third, we will hold 2 stakeholder co-design workshops involving eight clinicians and eight women with endometriosis separately and then bring all 16 participants together. Participants will discuss the implementation, delivery, usefulness, and sustainability of the prediction models. Clinicians will also focus on the ease of use and access to clinical prediction tools.

Results: This project was funded in March 2018 and approved by the Institutional Research Ethics Board in December 2019. At the time of writing, this study was in the data analysis phase, and the results are expected to be available in April 2021.

Conclusions: This study is the first to aim to predict who will benefit most from laparoscopic surgery through the reduction of pain or increased QoL. The models will provide clinicians with robustly developed and externally validated support tools, improving decision making in the diagnosis and treatment of women.

International Registered Report Identifier (IRRID): DERR1-10.2196/20986
endometriosis; algorithm; laparoscopy; pain; therapeutic

Introduction

Background on Endometriosis

Endometriosis is a chronic inflammatory condition affecting 6%-10% of women of reproductive age. It is defined by the presence of endometrial-like tissue outside the uterus (lesions), commonly affecting the pelvis and ovaries [1,2]. Although as common as back pain, it is poorly understood, and treatment and diagnosis are often delayed [2]. For example, in the United Kingdom, there is an average delay of 7-9 years in accessing treatment for endometriosis [3]. This is largely due to the lack of accurate, noninvasive diagnostic tests or biomarkers [4]. The diagnostic gold standard is pelvic laparoscopy under general anesthesia. Laparoscopy can be diagnostic, therapeutic, or both. Diagnostic delays lead to unnecessary suffering, and endometriosis is associated with debilitating pelvic pain, infertility, and fatigue and can have devastating effects on the quality of life (QoL). Endometriosis has no cure, but there are a number of treatment options. These include drugs that suppress ovarian function (which can have adverse effects [5]) or surgery for the lesions (usually laparoscopically). Surgical removal is often considered the best option for symptomatic endometriosis [6], but it does not reduce pain in 20%-28% of patients who undergo surgery [7,8].

Variable Response to Surgery

Quantifying the potential for improvement in a woman’s symptoms after surgery is important for guiding clinical decisions and treatment strategies. Secondary findings from observational, single-center studies indicate a graded response regarding pain reduction after endometriosis surgery, which is inversely related to disease severity [8-11]. One randomized controlled trial (RCT) found that pain symptoms improved after endometriosis surgery in significantly more patients with moderate and mild endometriosis (approximately 100% and 70%, respectively) than minimal disease (approximately 40%) [9]. In 2 other studies, women with deep endometriosis (DE) experienced more pain reduction after surgery than those with superficial endometriosis [10,11].

Although National Health Service (NHS) England recommends that women should undergo therapeutic laparoscopy for complicated DE in specialist endometriosis centers, we have no evidence-based information to recognize these women clinically. Nor are we able, at present, to recognize which women will respond to this treatment. As a result, therapeutic laparoscopy, a costly and limited resource with long waiting lists, is not necessarily carried out on those who will experience pain reduction.

Factors predicting pain reduction after endometriosis surgery have not yet been adequately identified, as they have never been studied as the primary research question. However, secondary outcomes from previous trials indicate that such factors can be identified [8,9,11]. Our aim is to address the existing gap by predicting success through pain reduction after endometriosis surgery.

Methods

The CRESCENDO (Creating a Clinical Prediction Model to predict Surgical Success in Endometriosis) project will be undertaken using existing recommendations for prognostic research model development, validation [12-14], and reporting [15]. It will also involve a systematic review of clinical risk factors associated with endometriosis and postsurgery treatment success, which will adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [16]. The systematic review is registered in the International Prospective Register of Systematic Reviews (PROSPERO).

Aim

To develop clinical prediction models for endometriosis to answer 2 fundamental questions regarding the surgical management of endometriosis:

1. Which women with confirmed endometriosis benefit from surgery and see improvement in pain and QoL? (Primary models)
2. Could these women be identified based on clinical symptoms measured before laparoscopy? (Secondary models)

Specific Objectives

1. To perform a systematic review of preoperative and intraoperative factors associated with postsurgical treatment success in endometriosis.
2. To perform a systematic review of the clinical risk factors associated with endometriosis.
3. To develop and validate clinical prediction models to predict changes in self-reported pain and QoL after surgery in women with a confirmed (primary models) or suspected (secondary models) diagnosis of endometriosis.
4. To describe an implementation plan within the secondary care pathway for women with a confirmed or suspected diagnosis of endometriosis using co-design workshops.

Systematic Review

We will carry out a systematic review and meta-analysis of observational cohort and case-control studies reporting 1 or more risk factors for endometriosis and predictors of postsurgical treatment success. We aim to determine the following:

- Absolute risk of having endometriosis in the presence or absence of a given risk factor.
- Relative risk of having endometriosis in the presence or absence of a given risk factor.
- The population attributable fraction for endometriosis in relation to each risk factor.
• Pre- and intraoperative factors for postsurgical treatment success for endometriosis and the associated risks.

**Literature Search**

We will search PubMed, Embase, and Cochrane databases from inception without any language restrictions and supplement these with manual searches of reference lists of included primary studies and relevant review articles.

**Study Selection, Data Extraction, and Quality Assessment**

The first reviewers will independently screen the titles and abstracts to identify the eligible studies, followed by retrieval and assessment of full texts of potentially relevant articles. Any disagreement will be resolved following discussions with a different reviewer. We will extract data in duplicate using predesigned data extraction forms. We will assess the quality of methodology of the included studies using the Newcastle Ottawa Scale [17] or the Jadad score [18], depending on the study.

**Data Sets for Development of the Clinical Prediction Models**

We have identified 3 suitable data sets that will be employed to develop and validate the clinical prediction models for women with confirmed and suspected diagnoses of endometriosis. We have guaranteed access to all three databases and data sharing agreements have been finalized *a priori*.

The British Society of Gynaecological Endoscopy (BSGE) maintains a national database that contains data of over 5000 women who had a confirmed diagnosis and underwent laparoscopic surgery for advanced endometriosis (stage 4). Records have been collected by clinicians at over 50 endometriosis centers in the United Kingdom since 2007 (data collection is ongoing). Endometriosis centers are commissioned by NHS England and accredited by the BSGE for complex multidisciplinary surgery required for the treatment of DE [19,20]. To maintain accreditation, the database serves as a mandatory record of DE cases and outcomes [21]. It includes data on patient characteristics, pain, and QoL before and after surgery, along with intraoperative findings. It is the most comprehensive source of data on endometriosis surgery and QoL worldwide. Some of these data have recently been published [22].

The second data set comes from a clinical study of women with suspected endometriosis, specifically chronic pelvic pain. *MRI versus laparoscopy to diagnose the main causes of chronic pelvic pain in women: a test-accuracy study and economic evaluation* (MRI to establish diagnosis against laparoscopy [MEDAL]) [ISRCTN13028601], was a comparative test-accuracy study assessing whether magnetic resonance imaging (MRI) could replace or triage the use of laparoscopy in establishing a diagnosis among women presenting in secondary care with chronic pelvic pain. Data were collected on patient characteristics, pain, and QoL before and after diagnostic laparoscopy with or without surgery, along with intraoperative findings for over 300 women who underwent laparoscopic surgery for chronic pelvic pain at 26 UK hospitals.

During surgery, over a third of the women were diagnosed with endometriosis [23].

The third data set included women with suspected endometriosis collected during an RCT. This study, *Laparoscopic Uterosacral Nerve Ablation (LUNA) for alleviating chronic pelvic pain* [ISRCTN41196151], randomized 487 women with chronic pelvic pain from 18 UK hospitals to assess the effectiveness of laparoscopic uterine nerve ablation. No significant improvements were reported on the visual analog pain scales. The data collected were similar to MEDAL and included patient characteristics, pain, and QoL at multiple time points [24].

**Establishment of Data Sharing**

The Queen Mary University of London pragmatic clinical trials unit (PCTU) will provide data management for secure data set transfer and storage in accordance with general data protection regulation and information governance principles. Data will be stored within a PCTU safe haven. No additional data collection will be needed. The MEDAL, LUNA, and BSGE data will be given to the PCTU in a pseudonymized form. The minimum data to be collected from the 3 data sets will be agreed upon by the study team, collaborators, and study steering committee. All recorded variables will be considered for collection.

**Data Preparation**

Data will be supplied in the format convenient for the original researchers. This project will take responsibility for converting, cleaning, and formatting the data as required before analysis. MEDAL and LUNA are data sets from previous funded studies and have been quality checked, analyzed, and published in peer-reviewed journals. Therefore, we will not assess the quality of the data, and we expect a limited need to clean them. Analyses of the BSGE data set have recently been published [22]. However, we will receive the raw source data, which will require cleaning, data quality checks, and assessment of the availability of relevant data for inclusion in the analysis. The collection of BSGE data is ongoing, so matching results from previous publications will be limited.

The 3 data sets (BSGE, MEDAL, and LUNA) will be employed when creating the prediction models, either as *development* or *external validation* data sets. To enable external validation of the prediction models, the predictors in the development data set will be matched with the variables in the validation data set. Where a direct match is not available in the data, we will investigate whether a new variable can be created from other information, such as calculating BMI from weight and height or categorizing continuous variables into groups.

**Outcomes**

Treatment success will be defined by changes in self-reported pain scores or QoL from baseline to 6 months or 1 year after surgery. Three months was considered too short, because there could be a placebo effect from the surgery [25]. One year was chosen as optimal for ensuring that postoperative healing was complete, and periods had returned (some women are given medication preoperatively to stop periods); the 6-month follow-up was included for pragmatic reasons as it is the longest follow-up duration many studies achieve, as ascertained in initial
scoping work. All three data sets have collected outcomes at the time points that enable this. Self-reported pain has been recorded in the data sets on a visual analog scale (VAS; score 0-10) and includes a range of specific pain symptoms such as dysmenorrhea (painful periods) and less specific symptoms such as chronic, noncyclical pelvic pain. QoL will be assessed using the EuroQol-5 Dimension questionnaire and a VAS on the overall health state (score 0-100).

After extensive discussions between the study team and the patient and public involvement (PPI) group, as well as interested clinicians, the most clinically relevant outcomes, to be determined from the data on women with menstrual cycles but also potentially relevant for those without them, have been chosen to be as follows:

1. Pain-dysmenorrhea
2. Pain-dyspareunia
3. Pain-chronic pelvic pain
4. Pain-dyschezia
5. QoL-overall health state

The PPI group and clinicians have also indicated a strong preference for a separate prediction model for each outcome instead of using a composite of multiple pain measurements. This will allow clinicians and patients to predict treatment success for a patient’s specific pain profile. All outcomes will be predicted on a continuous scale, rather than dichotomizing the change in score using arbitrary cut-offs.

In the MEDAL database, symptom duration is defined as the average level of pain over the last month. The BSGE database also records women’s self-rating of their pain over the last cycle. As these data are used to build the models, we suggest that the pain types that are input into the model are assessed over the same timescale. Women in the MEDAL study were followed up for 6 months. Hence, the computed output of change in pain is necessary for the timepoint of 6 months.

**Candidate Factors**

A list of candidate factors will be finalized before model development begins. Different factors might be considered in the analysis of treatment success in women with confirmed and suspected endometriosis.

Candidate factors will be identified through expert clinical input as well as through a systematic review. We will use the single factors that were investigated in narrative systematic reviews [19,26] and factors that were used as a part of previous models. These are likely to overlap. We will confirm that any candidate factor is available in the data for the analysis.

**Sample Size Considerations**

This study uses pre-existing data sets to develop and validate multivariable prediction models, and at the outset of this project, no formal guidance on the minimum sample size was available. The available records from all 3 data sets will be used. However, for confirmation and after completing the analysis, we will compare our results with recently published recommendations for sample size calculations in prognostic studies [27].

The BSGE data set contains records of approximately 5000 women, of which approximately one-third are complete. Data availability for analysis is likely to increase with the selection of candidate factors from the list of available variables.

The MEDAL data set contains records of over 300 women, with approximately 110 confirmed diagnoses of endometriosis. We assume no impact of the additional diagnostic test (MRI) performed in one arm of the study and therefore include the full study population in our analyses. The analysis of treatment success in women with suspected endometriosis will examine a reduced set of candidate factors that are available before diagnostic laparoscopy. We will ensure that the list of factors considered is appropriate for the available sample size.

The LUNA data set will serve as an external validation set for the model developed to predict treatment success in women with suspected endometriosis. We will include both the treatment and control arms of the trial, as no evidence of an effect of the LUNA intervention was found for any of the pain outcomes included. The LUNA data set contains records of over 590 women, with approximately 140 women with confirmed endometriosis.

**Model Development**

A total of 2 groups of models will be developed, reflecting the 2 different populations of women we aim to study. The group of 5 primary models will predict treatment success in women with confirmed diagnoses according to the 5 outcomes described above. The second group will consist of 5 models that predict treatment success in women with suspected endometriosis. Overall, this will result in 10 distinct models.

For the primary models (for women with a confirmed diagnosis of endometriosis), we will use the BSGE data set; a randomly selected 10% of records will be removed for performance testing of the model. For the secondary models of women with suspected diagnoses, we will use the MEDAL data set. All models will be developed using a logistic regression approach, linking candidate factors to outcomes.

A backward selection process will be used to decide which of the candidate predictor variables should be included in the final models (with a cutoff value of $P<.15$ conservatively taken to warrant inclusion and prevent omission of important predictors); candidate factors will likely vary between the models.

Continuous variables will be kept as continuous in the models (and not dichotomized) to avoid loss of power. Nonlinear effects may be considered if allowed by the sample size. No imputation of missing data will be performed, but the missing data at random assumption will be investigated.

**Model Performance and Internal Validation**

After developing the models, we will assess how well they perform. Calibration will be assessed visually using scatter plots; the calibration slope will indicate whether predictions are systematically too high or low (calibration-in-the-large). The apparent performance of the final models will be evaluated in terms of discrimination using the C statistic.
As apparent performance is often optimistic (due to a model being developed and validated in the same data set), internal validation, which we refer to as model performance, will also be undertaken. For the primary models of treatment success in women with confirmed endometriosis, we will randomly divide the BSGE records into 2 data sets, development (90% of the data) and validation (10% of the data). The apparent performance of the models in the validation sample will be compared with their performance in the development data set. Optimism is the difference between the apparent value in the validation sample and the observed value in the development data set. This optimism estimate is then subtracted from the model’s apparent performance to obtain an optimism-adjusted estimate of each measure of performance for the model.

The sample size in the MEDAL data set is not sufficient to use the same approach for the secondary models of treatment success in women with suspected endometriosis. If deemed appropriate, we may implement a bootstrap resampling technique whereby the apparent performance of the developed model in bootstrap samples is compared with its performance in the model developed using the original data set.

**External Validation and Recalibration**

In the last step, we will assess how well our models work when transferred into another setting (using another database).

The models of treatment success in women with confirmed endometriosis (ie, those developed in BSGE) will be externally validated in the MEDAL data set, and we will validate the models of treatment success in women with suspected endometriosis (ie, the ones developed in MEDAL) in the LUNA data set.

We will plot the agreement between observed and predicted change scores and assess calibration-in-the-large across deciles. In terms of discrimination, we will calculate the C statistic and its CI. We may consider updating the models, if they show poor performance in adjusting to the new situation, by carrying out recalibration or revision depending on discrimination and calibration performance. Table 1 shows how each data set will be used during these steps.

**Table 1. CRESCENDO (Creating a Clinical Prediction Model to predict Surgical Success in Endometriosis) data sets and their planned use during analysis**

<table>
<thead>
<tr>
<th>Analysis step and primary models: treatment success in women with confirmed diagnosis</th>
<th>Secondary models: treatment success in women with suspected endometriosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model development</td>
<td>Model performance</td>
</tr>
<tr>
<td>Random 90% of BSGE(^b) data set</td>
<td>Remaining 10% of BSGE data set</td>
</tr>
<tr>
<td>MEDAL(^c) data set</td>
<td>Bootstrapped samples of MEDAL data set or another appropriate method</td>
</tr>
<tr>
<td>External validation</td>
<td>Laparoscopic uterosacral nerve ablation (control arm only)</td>
</tr>
</tbody>
</table>

\(^a\)Databases containing pre-, intra-, and postoperative information of women with deep endometriosis (British Society of Gynaecological Endoscopy) or absent or superficial endometriosis (magnetic resonance imaging to establish diagnosis against laparoscopy or laparoscopic uterosacral nerve ablation).

\(^b\)BSGE: British Society of Gynaecological Endoscopy.

\(^c\)MEDAL: magnetic resonance imaging to establish diagnosis against laparoscopy.

**Presentation of the Prediction Models**

The final step of model development will be the translation of the models into easy-to-calculate scores. These scores will be presented against observed change scores, and if appropriate, they will be grouped by categories of treatment success. These categories may be defined in line with usual practice (eg, improvement by at least 1 point on VAS or change by less than 1 point on VAS or deterioration by at least 1 point on VAS). Descriptive comparisons will be presented for women in different categories of treatment success.

**Stakeholder Co-design Workshops**

Facilitated stakeholder co-design workshop discussions will center around a video made by the study team beforehand. The video will comprise 4 to 5 women, recruited as part of the study process, recounting their endometriosis treatment decision-making experience. Participants will be asked to identify and discuss key touch-points, that is, points within the endometriosis pathway where our clinical prediction model might have an emotional or clinical impact or where there may be impediments to its sustainable implementation. We will conceal patient identities in the video if they require this, for example, by masking their faces or having a member of the study team recount their experience on the video from a transcription.

A total of three workshops will be held. The first will involve 8 clinicians; in the second, we will work with 8 women with endometriosis; and in the third, we will bring all 16 participants together. Each workshop will last 2 h and all are planned to be on the same day, with the first 2 being simultaneous. We currently envisage the clinical predictor model to be on a computer screen during the sessions in the format we expect it to be used clinically. This proposed approach to using our model will be explored in the workshop and discussed in the context of the touch-point work, to determine what the potential issues, benefits, obstacles, and enablers are to the implementation, delivery, usefulness, and sustainability of this approach. We will ask the clinicians in the workshop to comment on the ease...
of use of the format and how they would like to access it (eg, an interactive formula on the BSGE website or embedded in guidelines). We will ask participants how the model should be used in secondary care consultations. We will also ask participants what is acceptable as a meaningful minimal change in pain score after surgery.

We will also ask the group to propose alternative approaches and solutions for any identified issues. For example, in our PPI work, it was suggested that the algorithm might be incorporated into a menstrual tracker app. As a follow-up from our study, we may develop a clinical trial of our clinical predictor model that builds on the workshop recommendations. This would need to consider the full patient pathway from primary care. Future trial considerations and the full patient pathway from primary to secondary care will therefore form a part of our workshop discussions.

Ethics Approval and Consent to Participate
The main project involves the analysis of anonymized data sets, and thus does not require ethical review. We will need an ethical review for the workshop. As this will be held at the end of the study and no other processes are dependent on it, the ethical review application will be prepared after the project has commenced.

Results
This project was funded in March 2018, approved by the Institutional Research Ethics Board in December 2019, and was in the phase of data analysis at the time the final revisions of this paper were made, with the results expected to be available in April 2021.

Discussion
Principal Strengths
This protocol defines the methods that will be applied to develop and externally validate our clinical prediction models to predict which women will benefit most from laparoscopic surgery resulting in reduction of pain or increased QoL. Previous models have focused on directly predicting endometriosis in patients with chronic pelvic pain and other symptoms, with limited success, or restricted their models to limited patient populations, with limited generalizability.

Our approach will address some of the challenges that other researchers have faced when attempting to improve care for women with chronic pelvic pain and endometriosis. As the prediction models focus on patient-perceived outcomes (QoL and pain), they will be more clinically relevant to the patient, and we are not limited by the wide range of definitions and treatments specific to endometriosis diagnosis, which often prevents external validation. The use of pre-existing data allows a comparatively quick and efficient development of the model with no unknown quantities, such as attrition and data completeness.

Limitations
The use of pre-existing data collected as part of other projects means that we will be limited by the data as recorded and will have no input as to how and what information is collected. Nonetheless, we have been able to assess the extent of missing data and found it to be sufficiently low. Power considerations have been based on truly available data, and the large sample size, specifically in the BSGE data set, will allow us to investigate a wide range of prognostic factors.

A second limitation is the matching of variables in the development and external validation data sets. This challenge is common to many studies that use pre-existing external validation data sets. We have been comparing how data were collected in the data sets and found most variables to be compatible; however, not all factors that we may identify as prognostic will also be available in the validation data set, thereby limiting the list of prognostic factors that can be validated externally.

Main Outputs and Access
This study will provide robustly developed and externally validated prediction models for postsurgery treatment success in women with suspected or confirmed endometriosis. The models will be generalized to a large range of women with this condition. To our knowledge, this project will be the first to predict who will benefit most from laparoscopic surgery resulting in reduction of pain or increased QoL and is therefore much needed. The prediction models will provide clinicians with a supporting tool for improving decision making in the diagnosis and treatment of women, thereby reducing unnecessary costs and harms associated with laparoscopic surgery.

Upon completion of the entire study, the models will be put in the public domain and will potentially be available for immediate use. The plan is that patient characteristics and clinical data can be entered into the formula by the user to calculate an individualized prediction of improvement after surgery. Presentation, implementation, and uptake within secondary care will be refined and finalized after the co-design workshops and beyond the lifespan of this project. For example, access could be in the form of a website or phone app. Our work with patient groups in the co-design workshops will give us invaluable direction on how to best advertise and deploy the prediction models, ensuring the greatest possible gain for patients and clinicians.

Acknowledgments
The authors would like to acknowledge all researchers, research nurses, and other staff involved in collecting, managing, and sharing the data used in this analysis, including Mel and Chloe.
The data to support the findings of this project are available from the PCTU data sharing committee, but restrictions apply to the availability of these data, which are used under license for this study and so are not publicly available. However, data are available from the authors upon reasonable request, with permission from the original data providers and a review by the data sharing committee.

Funding was received from the National Institute for Health Research-Research for Patient Benefit program.

Authors' Contributions

EB, EC, and AH developed the concept of this study. All authors participated in face-to-face meetings for deciding the study’s objectives and design aspects, such as definition of outcomes, data collection, and analysis strategy. All authors contributed significantly to the writing and review of the funding application. NM drafted the protocol manuscript with input on the clinical aspects from EB, EC, and AH. JA provided the systematic review section. Further invaluable contributions to the protocol were provided by CR who designed the qualitative part of the study and drafted the relevant section, and JA and JD, who refined and reviewed the conduct and management of the project. NM developed the analytical strategy. All authors have critically reviewed all subsequent versions of the manuscript and approved the final manuscript.

Conflicts of Interest

CR and AH have received honoraria for consultancy for Ferring, Roche, Nordic Pharma, and Abbvie. All other authors have no conflicts to declare.

References


Abbreviations
- **BSGE**: British Society of Gynaecological Endoscopy
- **DE**: deep endometriosis
- **LUNA**: laparoscopic uterosacral nerve ablation
- **MEDAL**: magnetic resonance imaging to establish diagnosis against laparoscopy
- **MRI**: magnetic resonance imaging
- **NHS**: National Health Service
- **PCTU**: pragmatic clinical trials unit
- **PPI**: patient and public involvement
- **PRISMA**: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
- **QoL**: quality of life
- **RCT**: randomized controlled trial
- **VAS**: visual analog scale
An Intervention for the Transition From Pediatric or Adolescent to Adult-Oriented HIV Care: Protocol for the Development and Pilot Implementation of iTransition

Amanda E Tanner, MPH, PhD; Nadia Dowshen, MSPH, MD; Morgan M Philbin, MHS, PhD; Kelly L Rulison, MS, MAS, PhD; Andres Camacho-Gonzalez, MD; Susan Lee, MPH; Shamia J Moore, MPH; J Dennis Fortenberry, MS, MD; Sophia A Hussen, MPH, MD

1Department of Public Health Education, University of North Carolina Greensboro, Greensboro, NC, United States
2Craig-Dalsimer Division of Adolescent Medicine, Children’s Hospital of Philadelphia, Philadelphia, PA, United States
3Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
4Department of Sociomedical Sciences, Columbia University Mailman School of Public Health, New York, NY, United States
5Department of Human Development and Family Studies, The Pennsylvania State University, State College, PA, United States
6Department of Pediatrics, Division of Infectious Diseases, Emory University School of Medicine, Atlanta, GA, United States
7Hubert Department of Global Health, Emory University Rollins School of Public Health, Atlanta, GA, United States
8Division of Adolescent Medicine, Indiana University School of Medicine, Indianapolis, IN, United States
9Department of Medicine, Division of Infectious Diseases, Emory University School of Medicine, Atlanta, GA, United States

Corresponding Author:
Amanda E Tanner, MPH, PhD
Department of Public Health Education
University of North Carolina Greensboro
437 Coleman Building
Greensboro, NC, 27402-6169
United States
Phone: 1 715 497 2885
Email: aetanner@uncg.edu

Abstract

Background: In the United States, adolescents and young adults are disproportionately affected by HIV and have poorer HIV-related health outcomes than adults. Health care transition (HCT) from pediatric or adolescent to adult-oriented HIV care is associated with disruptions to youths’ care retention, medication adherence, and viral suppression. However, no evidence-based interventions exist to improve HCT outcomes for youth living with HIV.

Objective: There are 2 phases of this project. Phase 1 involves the iterative development and usability testing of a Social Cognitive Theory–based mobile health (mHealth) HIV HCT intervention (iTransition). In phase 2, we will conduct a pilot implementation trial to assess iTransition’s feasibility and acceptability and to establish preliminary efficacy among youth and provider participants.

Methods: The iterative phase 1 development process will involve in-person and virtual meetings and a design team comprising youth living with HIV and health care providers. The design team will both inform the content and provide feedback on the look, feel, and process of the iTransition intervention. In phase 2, we will recruit 100 transition-eligible youth across two clinical sites in Atlanta, Georgia, and Philadelphia, Pennsylvania, to participate in the historical control group (n=50; data collection only) or the intervention group (n=50) in a pilot implementation trial. We will also recruit 28 provider participants across the pediatric or adolescent and adult clinics at the two sites. Data collection will include electronic medical chart abstraction for clinical outcomes as well as surveys and interviews related to demographic and behavioral characteristics; Social Cognitive Theory constructs; and intervention feasibility, acceptability, and use. Analyses will compare historical control and intervention groups in terms of HCT outcomes, including adult care linkage (primary), care retention, and viral suppression (secondary). Interview data will be analyzed using content analysis to understand the experience with use and acceptability.
Results: Phase 1 (development) of iTransition research activities began in November 2019 and is ongoing. The data collection for the phase 2 pilot implementation trial is expected to be completed in January 2023. Final results are anticipated in summer 2023.

Conclusions: The development and pilot implementation trial of the iTransition intervention will fill an important gap in understanding the role of mHealth interventions to support HCT outcomes for youth living with HIV.

International Registered Report Identifier (IRRID): DERR1-10.2196/24565

(KEYWORDS
HIV; mHealth; transition to adult care; young adult; feasibility studies; retention in care; control groups; United States; telemedicine; HIV infections; mobile phone

Introduction

Adolescents and young adults are disproportionately affected by HIV in the United States compared with their adult counterparts [1]. Youth living with HIV aged 18 to 24 years face significant barriers to continuous care engagement. Recent research suggests that only 43% of youth living with HIV have received any HIV medical care and that only 33% are virally suppressed [1]. These challenges are accentuated during the health care transition (HCT) from pediatric or adolescent to adult-oriented care [2,3]. Every youth living with HIV in pediatric or adolescent care must transition to adult care, and 25,000 youth living with HIV in the United States are projected to undergo HCT in the decade leading up to 2025 [4]. In the United States, HIV-related HCT typically occurs around the age of 24 years [5] and is often associated with disruptions to care retention, medication adherence, and viral suppression [6]. In the 1 to 2 years following expected HCT, less than 40% of youth are linked to adult care and only 50% to 60% of those linked remain engaged in care [6-9]. A seamless HCT experience can facilitate continuous care retention and viral suppression in youth living with HIV, with implications for morbidity, mortality, and public health [10]. Although multilevel barriers to HCT exist [11], existing research focuses primarily on single-site protocols that have not been rigorously evaluated, and as such, no evidence-based interventions exist to improve HCT outcomes for youth living with HIV [4,5].

Mobile health (mHealth) interventions are used to improve medication adherence, care retention, and viral suppression among youth living with HIV [12-15]. mHealth interventions are well suited for youth living with HIV: youth use the internet more than any other age group, and over 90% own smartphones [16]. In addition, providers frequently use mobile platforms in health care settings to facilitate provider-provider and patient-provider communication and to standardize clinical practices [17-19]. Therefore, an mHealth platform is an ideal vehicle for a novel multilevel intervention to improve HCT, with a high likelihood of adoption by youth living with HIV and their pediatric or adolescent and adult providers.

Theory-informed mHealth interventions can improve health outcomes for youth living with HIV, and Social Cognitive Theory [20] is one of the most widely used theoretical frameworks in mHealth behavioral interventions [21,22]. Social Cognitive Theory helps to articulate dynamic interactions that occur during HCT—between environment (clinical environments), behavior (care engagement through HCT), and personal factors (including self-efficacy for facilitating HCT and knowledge of the HCT process)—and is helpful in understanding the multilevel challenges associated with HCT. Accordingly, this paper describes the development and evaluation protocol for iTransition, a Social Cognitive Theory–based mHealth intervention to improve HCT processes for youth living with HIV and their pediatric or adolescent and adult care providers.

Methods

Ethics Statement

The Institutional Review Board (IRB) at the Children’s Hospital of Philadelphia (CHOP) is the IRB of record for all the participating institutions (CHOP, Emory University, and the University of Carolina Greensboro). The CHOP IRB has reviewed and approved all the procedures outlined in this protocol. The study is registered with ClinicalTrials.gov (NCT04383223).

Study Settings

This protocol is being implemented at high-volume HIV care centers that frequently transition youth living with HIV from pediatric or adolescent to adult-oriented care. The Grady Infectious Disease Program in Atlanta, Georgia, is affiliated with a large public safety net health system and is among the largest HIV care centers in the United States, serving approximately 6000 patients per year, 10% of whom are youth living with HIV under 25 years of age. It contains pediatric or adolescent and adult-oriented clinic spaces in the same building but has no formal HCT protocol. CHOP, located in Philadelphia, Pennsylvania, contains 2 academic hospital-based HIV clinics that transition youth living with HIV to multiple adult clinics, including infectious disease clinics at the Hospital of the University of Pennsylvania and the Penn Presbyterian Medical Center. CHOP has a formal HCT protocol that includes guided pre-HCT visits to adult clinics with a pediatric or adolescent social worker. Pilotting iTransition in clinics in Atlanta and Philadelphia with different HCT models (eg, one colocated vs multiple adult clinic options) enhances the likelihood of generalizability to other HIV clinical care settings.
Intervention Overview

*iTransition* is grounded in Social Cognitive Theory and builds on our team’s extensive HCT work [6,7,11,23,24]. Broadly, Social Cognitive Theory outlines the dynamic interplay, or *reciprocal determinism*, between environmental, behavioral, and personal factors. Specifically, Social Cognitive Theory posits that these factors interact through constructs, including self-management skills, outcome expectancy, self-efficacy, cues to action, and behaviors, each of which forms intervention targets and situates our approach to improving HCT for youth living with HIV. Our previous research outlined 4 domains of successful HCT [6] that align with these constructs and will be addressed in *iTransition* (Table 1): (1) preparation of youth living with HIV for autonomous disease management, (2) effective interclinician provider communication, (3) enhancement of youth living with HIV’s connectedness to adult clinics, and (4) implementation of formalized HCT protocols in pediatric or adolescent and adult clinics. Intervention content and strategies will include educational modules, interactive activities (eg, quizzes), relationship development strategies (eg, youth and provider profiles), and communication (eg, individual and group chat features).

In addition, we draw on supportive accountability theory [25], which posits that human interactions with coaches—here, providers serving as Transition Champions—provide an essential complement to mHealth interventions that enhance uptake for users. In phase 2, Transition Champions will assist other participants with using the *iTransition* app, encourage utilization by their peers (providers) and patients (youth), and promote integration of *iTransition* use into clinical workflow.

<table>
<thead>
<tr>
<th>Domains of HCT(^a) success</th>
<th>Measurable social cognitive theory constructs</th>
<th><em>iTransition</em> intervention strategies (target audience)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation of YLH(^b) for autonomous disease management</td>
<td>• Self-management skills&lt;br&gt;• Self-efficacy</td>
<td>• Youth-friendly educational modules and readiness assessments (youth)&lt;br&gt;• HCT educations modules (providers)&lt;br&gt;• Interactive case scenarios (youth and providers)</td>
</tr>
<tr>
<td>Enhancing YLH’s connectedness to adult clinics</td>
<td>• Self-management skills&lt;br&gt;• Self-efficacy</td>
<td>• Motivational messages (youth)&lt;br&gt;• TC(^c) interactions (providers)&lt;br&gt;• Online support forum (youth)</td>
</tr>
<tr>
<td>Effective between clinic provider communication</td>
<td>• Self-efficacy&lt;br&gt;• Behavior</td>
<td>• Reminders and updates about patients going through HCT process (providers)&lt;br&gt;• Communication between adult and pediatric or adolescent providers (providers)</td>
</tr>
<tr>
<td>Implementation of formalized HCT protocols in pediatric or adolescent and adult clinics</td>
<td>• Behavior&lt;br&gt;• Cues to action</td>
<td>• Clinic-level implementation (youth)&lt;br&gt;• TC intervention (providers)</td>
</tr>
</tbody>
</table>

\(\text{aHCT: health care transition.}\)<br>
\(\text{bYLH: youth living with HIV.}\)<br>
\(\text{cTC: Transition Champion.}\)

Study Design

Phase 1 of this study, *iTransition* intervention development, uses a participatory, theoretically driven approach. In phase 2, we will conduct a pilot implementation trial of *iTransition* in Atlanta and Philadelphia, comparing post-HCT clinical HIV outcomes—adult care linkage (primary), care retention, and viral suppression (secondary)—between youth participants in our intervention with historical control groups. See Multimedia Appendix 1 for *iTransition* screenshots.

Phase 1: Intervention Development

To develop *iTransition*, we will work collaboratively and iteratively with a design team that includes pre- and post-HCT youth (n=9) and pediatric or adolescent and adult-oriented HIV care providers (n=8) in Atlanta and Philadelphia (Figure 1). The provider participants in the design team will also serve as Transition Champions during the phase 2 implementation trial. On the basis of the input from the design team, we will work with our technology partner, Pattern Health, to create intervention content in the youth-facing mobile application (ie, *app*) and the provider-facing console. We will present ideas (including app and console demonstrations) to the design team members to solicit feedback for revisions on content, format, and presentation until the app is completed and ready for usability testing. This process will include regular in-person and virtual design team meetings. Once the completed mHealth intervention is built, we will conduct usability testing with at least five youth living with HIV and five providers to get further feedback on the functionality and app use experience. Specifically, these youth and providers will be asked to use *iTransition* and respond to a series of assessment questions aligned with several of the user experience honeycomb constructs—usable, useful, valuable, desirable, and credible [26,27].
Phase 2: Pilot Implementation Trials

In phase 2, we will conduct a pilot implementation trial of iTransition with 100 transition-eligible (ie, within 6 months of anticipated HCT) youth living with HIV recruited by staff in participating pediatric or adolescent HIV clinics and 28 providers (inclusive of the 8 Transition Champions recruited in phase 1) from pediatric or adolescent and adult HIV clinics (Figure 2). Due to the challenges of randomizing involvement in a clinic-wide intervention, we will compare our intervention group with a historical control group that will be completely recruited and enrolled before implementation of the iTransition intervention. We will recruit the historical control group (n=50) from both study sites (approximately 25 per site). Following the recruitment of our historical control group, we will recruit intervention youth (n=50; approximately 25 per site) and the remaining 20 providers. We will compare the youth historical control and intervention groups in terms of their HCT outcomes. Our primary outcome is linkage to adult care, defined as attending one HIV care visit at the adult clinic. Our secondary outcomes are care engagement at the adult clinic (ie, at least two visits within 12 months, at least 90 days apart) and viral suppression (defined as less than 200 copies/mL) 1 year after baseline. We will also assess the use, acceptability, and feasibility of iTransition among youth and the provider and Transition Champion participants.
Figure 2. Phase 2 activities by participant type.

**Aim 2: iTransition Pilot (N=128)**

**Year 1:**
- **YLH historical controls (n=50)**
  - Single visit assessment
  - REDCap surveys
  - Interview 1
- **YLH intervention (n=50)**
  - 6-month baseline visit
  - Transition orientation
  - REDCap surveys
- **Youth Historical Control Group**
  - Inclusion criteria include (1) living with HIV, (2) age 18+ years, (3) expected to undergo HCT within the next 6 months, and (4) enrolled in care at participating clinics in Atlanta and Philadelphia. Potentially eligible youth will be identified through chart review in collaboration with clinic staff and approached by a research assistant during or after regular HIV care visits at the pediatric or adolescent clinic (face-to-face or during telehealth visits). Those who express interest will be consented (including signing a consent form for the release of medical information) and screened to determine eligibility. Those who enroll will complete a single cross-sectional survey before HCT, and they will receive US $25 for their time.

**Year 2:**
- **YLH intervention (n=50)**
  - 6-month follow-up visit
  - Functionality check REDCap surveys
- **Youth Intervention Group**
  - Following the recruitment and participation of the historical control group, we will recruit youth living with HIV for the intervention group. Inclusion criteria for the intervention group will be identical to the historical control group with the additional criteria that they must own a smartphone or tablet and report consistent internet access (defined as no lapse >24 hours in the last 6 months). Youth without a device are not eligible for this pilot trial; larger future trials will provide access to devices as needed. Similar to the historical control group, a research assistant will invite potentially eligible youth to participate during or after regular HIV care visits at the pediatric or adolescent clinic (face-to-face or during telehealth visits). Those who express interest will provide consent (including the release of medical information) and be screened to determine eligibility. Enrolled participants will complete the same baseline survey as the historical control group at the first study visit, during which time they will also be oriented to the iTransition app and create a user account. Over the 12-month intervention period, youth intervention participants will complete additional surveys at 6 and 12 months, and a subset (up to 30 youth) will be invited to participate in qualitative interviews based on iTransition use (high and low/no). Overall, participants could receive US $75 to US $125 (US $25 per study visit or interview) for their time if they complete all assessments (baseline, 6-month, and 12-month assessments and up to 2 qualitative interviews).

Participants

For phase 1, we will recruit approximately 10 youth and 8 providers for our design team. Phase 2 will include a total of 128 participants: 100 transition-eligible youth (50 for the intervention group and 50 for the historical control group), 20 providers (approximately 10 from each city site) who work with transitioning youth living with HIV, and the 8 design team providers (recruited in phase 1) who will function as Transition Champions provider participants. Historical control youth will participate for 1 day, intervention group youth will participate for 12 months, and all providers (including Transition Champions) will participate for 18 months.

Youth Historical Control Group

Inclusion criteria include (1) living with HIV, (2) age 18+ years, (3) expected to undergo HCT within the next 6 months, and (4) enrolled in care at participating clinics in Atlanta and Philadelphia. Potentially eligible youth will be identified through chart review in collaboration with clinic staff and approached by a research assistant during or after regular HIV care visits at the pediatric or adolescent clinic (face-to-face or during telehealth visits). Those who express interest will be consented (including signing a consent form for the release of medical information) and screened to determine eligibility. Enrolled participants will complete the same baseline survey as the historical control group at the first study visit, during which time they will also be oriented to the iTransition app and create a user account. Over the 12-month intervention period, youth intervention participants will complete additional surveys at 6 and 12 months, and a subset (up to 30 youth) will be invited to participate in qualitative interviews based on iTransition use (high and low/no). Overall, participants could receive US $75 to US $125 (US $25 per study visit or interview) for their time if they complete all assessments (baseline, 6-month, and 12-month assessments and up to 2 qualitative interviews).
Provider Intervention Group

Inclusion criteria for the provider intervention group include (1) staff members at participating clinics, (2) work with transitioning youth, and (3) access to the internet via any device (eg, a smartphone, tablet, or computer). Research assistants will directly invite providers at each of the participating pediatric or adolescent and adult clinics. Enrolled participants will complete a baseline survey and be oriented to iTransition and create a user account. During the 18-month intervention period, provider intervention participants will complete surveys at 6, 12, and 18 months, and a subset (up to 12) will be invited to participate in up to 3 qualitative interviews based on iTransition use (high and low/no). Overall, participants can receive up to US $100 to US $175 (US $25 per survey and qualitative interview) for their time over the course of the study.

Transition Champion Group

Inclusion criteria for the Transition Champion group include the same criteria as provider participants and nomination by a clinic staff member in their clinic to be an iTransition intervention point person (champion) who will support iTransition intervention use in their clinics. Following the baseline survey, orientation to iTransition, and a brief training, Transition Champions will complete surveys at 6, 12, and 18 months and be invited to participate in 3 qualitative interviews. Given the intensity and duration of their participation across the project, Transition Champions will receive a total of US $350 (US $50 for each study visit or interview) for their active support of iTransition use in their respective clinics by youth and provider groups.

Sample Size and Power

The goals of this study are to estimate effect sizes so that we can fully power a future randomized control trial and to evaluate iTransition acceptability and feasibility. Our youth sample size is sufficient to detect large differences in our primary HCT outcome (linkage to adult care) between those using iTransition compared with the historical controls. Our sample size of 50 subjects per group provides 87% power to detect a 30% difference in HCT rates between the historical control and intervention groups, 71% power to detect a 25% difference, and 52% power to detect a 20% difference. These calculations are based on an estimated HCT rate in the historical control group of 40% (over the 12-month period), as our previous work showed that 37% of youth living with HIV transitioned into adult-oriented care settings within a 9-month period [6].

Data Collection

Data collection for surveys and electronic medical record abstraction will occur in Research Electronic Data Capture [28]. Qualitative interviews will be digitally recorded, transcribed by a professional agency, and analyzed by the study team. Paradata (automatically collected electronic process data, eg, the number of times that a participant logged into the app) will be captured for all youth and provider intervention participants.

Youth Historical and Intervention Groups

Youth participants will complete a baseline survey with questions related to demographics, general health status, health behaviors (eg, substance use and medication adherence), Social Cognitive Theory–related factors, and other psychosocial covariates affecting HCT (eg, stigma, discrimination, social support, self-efficacy, skills, and readiness). The majority of measures have been validated and used in previous research (Multimedia Appendix 2 [29-44]). In line with the Social Cognitive Theory, we developed a measure to assess youth’s HCT-specific cues to action (eg, reminders to take medication and refill prescriptions, provider discussions related to HCT, provider-initiated skill development for adult clinic, and developing a written a HCT plan with the provider). We are also assessing incarceration history.

Intervention youth will also complete subsequent surveys at 6 and 12 months, including all the same baseline measures (except demographics). A subset of intervention youth will be invited to complete interviews related to iTransition use (eg, perceptions, experience, challenges, and recommendations for improving the intervention). Overall iTransition feasibility (ease of use), acceptability (tolerance of use), and use (eg, number of times in app and time with activities) data will be collected through survey and paradata.

We will review the electronic medical charts of youth in both study groups for a 3-year period: 12 months before baseline through 24 months after baseline. The use of medical chart data helps address potential attrition over time by allowing the assessment of outcomes (eg, medial appointment attendance and viral suppression). Chart reviews will be completed at 3 time points: baseline, 12 months after baseline, and 24 months after baseline. Each abstraction will capture data from the previous 12 months (ie, initial abstraction includes the 12 months before baseline, 12-month abstraction covers baseline to 12 months, and 24-month abstraction includes data between 12 months and 24 months). Information abstracted will include HIV history (eg, date of diagnosis, antiretroviral therapy information, and other sexually transmitted infection [STI] diagnoses), viral load, CD4+ count and percentage, diagnoses of STIs, and appointment information at pediatric or adolescent and adult clinics.

The primary clinical outcome variable for youth, measured at the patient level, is linkage to adult care (defined dichotomously as having 1 completed adult clinic appointment or not). Secondary clinical outcomes are care retention (dichotomously defined as having or not having 1 visit in each 6-month period) and viral suppression (<200 copies/mL) at 1 year after baseline. We will explore whether individual characteristics (eg, gender, race or ethnicity, and sexual orientation), Social Cognitive Theory–related covariates (eg, self-efficacy, and intervention dosage (as quantified by paradata metrics) are associated with outcomes or whether there is any preliminary evidence that it moderates intervention effects.

Provider and Transition Champion Groups

Data collection for the provider intervention and Transition Champion groups includes baseline and follow-up (at 6, 12, and 18 months) surveys (Multimedia Appendix 2). These will collect information related to demographics and professional experience (baseline only), clinic assessment (eg, leadership and self-efficacy to support HCT), and iTransition evaluation.
(Multimedia Appendix 2). Similar to the youth measures, the majority of provider measures are validated and have been used previously. The one measure that was not previously validated is related to general perceptions of the intervention (eg, barriers to intervention implementation, helpfulness of the intervention for supporting HCT, and motivations to use the intervention). A subset of providers and all of the Transition Champions will also be invited to complete interviews related to iTransition use (eg, perceptions, experiences, challenges, and recommendations for improving the intervention). Similar to youth intervention participants, overall iTransition use will be collected through survey and paradata. Primary outcomes for providers will be related to the use, feasibility, and acceptability of iTransition. In addition, changes in providers’ HCT-related self-efficacy will be examined.

Outcome Analyses

Quantitative Analyses: Youth Data
To estimate intervention efficacy, we will use a logistic regression model with a term for intervention status to compare the likelihood of being linked to adult care between the intervention and historical control groups. We will conduct univariable and multivariable logistic regression analyses to test the intervention effects on our secondary outcomes of care retention and viral suppression after 12 months. Retention for patients who never achieve HCT or who undergo HCT too late in the 12-month follow-up period to calculate the retention outcome will be treated as missing. Other outcomes will include acceptability and feasibility of the app, app use metrics, and Social Cognitive Theory–related measures, such as self-efficacy for managing HCT.

Quantitative Analyses: Provider and Transition Champion Data
We will conduct descriptive statistics to assess the feasibility, acceptability, and use of iTransition. We will also conduct regression analyses to assess changes in providers’ self-efficacy in facilitating HCT for their patients over time.

Qualitative Analyses: Youth, Provider, and Transition Champion Data
Our interview data will help describe participants’ experiences with iTransition and understand the processes of implementing the intervention in different clinical environments. Qualitative analyses will be guided by content analysis, which is well suited for understanding participants’ experiences and comparing experiences within and across groups (eg, comparing youth from Atlanta and Philadelphia or comparing pediatric or adolescent and adult providers). We will develop a preliminary codebook [45] to include predetermined deductive codes related to the Social Cognitive Theory domains of interest and inductive codes that emerge from the data. Transcripts will be coded by 2 members of the research team, and differences will be discussed and resolved in team meetings until consensus is reached. Data analysis will also involve generating frequencies of codes and comparing the frequency of code occurrence across participant subgroups. To further establish the trustworthiness and validity of our data, we will present findings back to our youth design team members (who have recently undergone HCT and who will participate in the development and usability testing of iTransition).

Results

Phase 1 (development) of iTransition research activities began in November 2019 and is ongoing. The data collection for the phase 2 pilot implementation trial is expected to be completed in January 2023. Final results are anticipated in summer 2023.

Discussion

Although every youth in pediatric or adolescent HIV care will need to transition to adult-oriented care, there are no evidence-based HCT interventions [4,6]. HCT poses a persistent challenge to the health of youth living with HIV as they may disengage from care resulting in gaps in medication adherence and viral rebound [7]. Thus, interventions to support youth and providers at both pediatric or adolescent and adult clinics are crucial. iTransition, as a theory-based, stakeholder-engaged, multilevel mHealth intervention, is particularly poised to fill this important gap in HCT for youth and providers.

Acknowledgments

This work was supported by the National Institutes of Mental Health (R34MH116805; principal investigators Hussen and Tanner). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors acknowledge the contribution of the research team in Atlanta and Philadelphia, including Kayla Knowles’ role in the development of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of iTransition.
[DOCX File, 3561 KB - resprot_v10i4e24565_app1.docx ]
Multimedia Appendix 2
Primary and other outcome measures: operationalization and schedule.

References


Abbreviations

CHOP: Children’s Hospital of Philadelphia
HCT: health care transition
IRB: Institutional Review Board
mHealth: mobile health
STI: sexually transmitted infection

©Amanda E Tanner, Nadia Dowshen, Morgan M Philbin, Kelly L Rulison, Andres Camacho-Gonzalez, Susan Lee, Shamia J Moore, J Dennis Fortenberry, Sophia A Hussen. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 07.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Technology-Based Fall Risk Assessments for Older Adults in Low-Income Settings: Protocol for a Cross-sectional Study

Ladda Thiamwong¹,², PhD; Jeffrey R Stout³,⁴, PhD; Joon-Hyuk Park⁴,⁵, PhD; Xin Yan⁵, PhD

¹College of Nursing, University of Central Florida, Orlando, FL, United States
²Disability, Aging and Technology Cluster, University of Central Florida, Orlando, FL, United States
³School of Kinesiology and Physical Therapy, College of Health Professions and Sciences, University of Central Florida, Orlando, FL, United States
⁴Mechanical and Aerospace Engineering Department, University of Central Florida, Orlando, FL, United States
⁵Department of Statistics and Data Science, University of Central Florida, Orlando, FL, United States

Abstract

Background: One-third of older adults have maladaptive fall risk appraisal (FRA), a condition in which there is a discrepancy between the level of fear of falling (FOF) and physiological fall risk (balance performance). Older adults who overestimate their physiological fall risk and report a high FOF are less likely to participate in physical activity. Limited data suggest that the association among FOF, body composition, and physical activity intensity differs by fear severity.

Objective: This study aims to examine the associations among FRA, body composition, and physical activity using assistive health technology, including the BTrackS balance system, bioelectrical impedance analysis, and activity monitoring devices. This study also aims to examine the feasibility of recruitment and acceptability of technologies and procedures for use among older adults in low-income settings.

Methods: This cross-sectional study will be conducted in older adults’ homes or apartments in low-income settings in Central Florida, United States. Following consent, participants will be contacted, and our team will visit them twice. The first visit includes questionnaire completion (e.g., sociodemographic or FOF) and balance performance test using the BTrackS balance system. The participants will be stratified by the FRA matrix. In addition, they will perform hand grip strength and dynamic balance performance tests. Participants will then be asked to wear the ActiGraph GT9X Link wireless activity monitor on the nondominant wrist for 7 consecutive days. The second visit includes body composition testing and a structured interview about the acceptability of the technologies and procedures.

Results: Ethical approval was obtained from the institutional review board of the University of Central Florida (protocol number 2189; September 10, 2020). As of December 2020, participation enrollment is ongoing and the results are expected to be published in Summer 2022.

Conclusions: Accurate FRA is essential for implementing physical activity programs, especially in older adults with low income. This study will provide data for developing technology-based fall risk assessments to improve participation in physical activity, thus enhancing healthy longevity among older adults in low-income settings.

International Registered Report Identifier (IRRID): PRR1-10.2196/27381

(JMIR Res Protoc 2021;10(4):e27381) doi:10.2196/27381

KEYWORDS

body composition; falls; risk assessment; technology; wearable devices; accidental falls; fear
**Introduction**

**Background**

More than 15 million (30%) older adults in the United States have income below 200% of the poverty line [1, 2]. The poverty rate increases with age and is higher among women, Black and Hispanic individuals, and individuals with poor health [1]. Older adults who live in low-income communities are less likely to engage in physical activity (PA) [3-5], defined as any bodily movement produced by skeletal muscles that results in energy expenditure [6]. Lack of PA is related to chronic conditions and reduced quality of life among older adults with low income [7, 8]. Limited data suggest that older adults who overestimate their fall risk and report fear of falling (FOF) are less likely to participate in PA, and the association between FOF and PA intensity differs by fear severity [9]. Changes in body composition also have a significant effect on physical functions and quality of life [10]. However, these studies did not include older adults from low-income settings, who were more likely to report falls in the previous year [11].

One-third of older adults have maladaptive fall risk appraisal (FRA), a condition in which there is a discrepancy between the perceived fall risk (levels of FOF) and physiological fall risk (balance performance) [12, 13]. Measuring FRA in older adults can be challenging because of self-report bias and cognitive deficit [14, 15]. Using both subjective and objective measures provides better fall risk assessment among older adults [15, 16]. Thus, we developed an FRA matrix, a graphical grid categorizing the levels of FOF and balance performance into 4 quadrants: (1) rational FRA (low FOF and normal balance), (2) incongruent FRA (low FOF despite poor balance), (3) irrational FRA (high FOF despite normal balance), and (4) congruent FRA (high FOF and poor balance) [12]. In our pilot study (n=102), we measured FOF using a questionnaire and balance performance using a portable and novel force balance plate (BTrackS Assess Balance System [BBS], Balance Tracking System, Inc). We found that 40% of older adults had maladaptive FRA. In this group, 19% of patients with incongruent FRA and 30% with irrational FRA reported falling in the past year [12]. Older adults with congruent FRA or high fear-high physiological were 2.14 times more likely to fall than those with rational FRA or low fear-low physiological [17]. However, our sample included only 11% of older adults with low income.

**Objectives**

Body composition (eg, obesity or low relative skeletal muscle mass) and decline in muscle strength have been associated with FOF, functional impairment, and disability in older American adults [18-20]. Higher BMI and percentage of body fat were associated with poor physical function, whereas the percentage of appendicular lean mass was associated with better physical function (eg, walking or balance test) [19]. Older adults with higher daily PA have better physical function than those who engage in less PA [21]. Although BMI is the most widely used measurement to classify overweight and obesity statuses, it is prone to measurement error and does not consider body fat distribution and skeletal muscle mass [22]. We will use the bioelectrical impedance analysis (BIA) device, which measures body composition and has established normative data among older adults [23].

Research has not examined the associations among body composition, FRA, and PA using assistive health technology (AHT), which applies organizational knowledge, skills, procedures, and systems to improve functioning [24]. This study examines the associations among body composition, FRA, and PA using AHT, including BIA [25], BBS [26, 27], and activity monitoring devices [28]. It also aims to examine the feasibility of recruitment and acceptability of technologies and procedures for use among older adults in low-income settings.

**Methods**

**Study Design**

This is a cross-sectional study.

**Settings**

This study will be conducted in older adults’ homes or apartments in low-income settings in Central Florida, United States. Collectively, approximately 300 older adults are available for screening at the planned sites.

**Study Participants**

We will enroll a sample of 120 participants, if they meet all of the following inclusion criteria: (1) aged ≥60 years, (2) have a low income (using poverty thresholds for 2019 by family size and the number of children aged <18 years, published by the US Census Bureau) [29], (3) have no marked cognitive impairment (memory impairment screen score ≥5 [30-32]), and (4) live in their own homes or apartments. Exclusion criteria were as follows: (1) a medical condition precluding balance test (eg, inability to stand on the balance plate) or PA (eg, shortness of breath or feeling pressure when performing PA), (2) currently receiving treatment from a rehabilitation facility, or (3) having medical implants (eg, pacemakers).

**Consideration of Sex and Other Relevant Biological Variables**

In our pilot study, 66% of the participants were women (mean 77, SD 7.6 years). We found that men tend to cluster in rational and incongruent FRA groups, whereas women were widely distributed among 4 groups (rational, incongruent, irrational, and congruent). We expect that 65% to 70% of our sample will be women. We will explore the differences between men and women and between age groups (<75 years vs >75 years) to assess whether sex and age as biological variables contribute to the findings.

**Variables and Instruments**

The study instruments consist of objective and subjective measures (Table 1).
Balance performance will be assessed using the BBS [26,27]. The BBS includes a portable BTrackS balance plate (BBP) and BTrackS Assess Balance software running on a computer device. The BBP dimensions are 15.5 × 23.5 × 2.5 in, weight 6.58 kg, and is operated on Windows 7 or higher via a USB port (Food and Drug Administration approved). During the test, a piece of sturdy furniture or a standard walker will be placed within the participant’s reach to reduce the risk of FOF contaminating the performance and to enable even frail people to participate [42]. The software uses the BBS normative database to compare the individual with others in the same age group. The BBS score is dependent on age and sex, but not body size, so that the percentile rankings can be determined across various age groups and for men and women separately [43]. A scale from 0 to 100 represents the percentile ranking of the BBS. A score of 0 to 30 indicates low fall risk (normal balance) and ≥31 indicates moderate-to-high fall risk (poor balance) [43]. The BBS has excellent validity using Pearson correlations (r>0.90) and high test-retest reliability (intraclass correlation coefficient [ICC]=0.83) [44].

Body composition will be assessed using a direct segmental multifrequency BIA using the InBodys10 device [25]. The BIA is manufactured by Biospace Corporation Limited. The BIA

### Table 1. Description of study variables and instruments.

<table>
<thead>
<tr>
<th>Variable and instrument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participant’s characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Sociodemographic and medical history</td>
<td>Self-report questionnaire</td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td>Patient Health Questionnaire-9 [33,34]</td>
</tr>
<tr>
<td>Cognition</td>
<td>MIS[^a] [30-32]</td>
</tr>
<tr>
<td>FOF[^b]</td>
<td>Short FES-F [35]</td>
</tr>
<tr>
<td>Balance performance</td>
<td>BTrackS balance system [26,27]</td>
</tr>
<tr>
<td>Balance performance</td>
<td>Sit-to-stand test [40,41]</td>
</tr>
<tr>
<td>Hand grip strength</td>
<td>A hand grip dynamometer (JAMAR 5030J1)</td>
</tr>
<tr>
<td>Body composition</td>
<td>Bioelectrical impedance analysis: InBodys10 device [25]</td>
</tr>
<tr>
<td>Physical activity</td>
<td>ActiGraph GT9X Link wireless activity monitors (ActiGraph LLC) [28]</td>
</tr>
<tr>
<td>Acceptability of technology</td>
<td>Evaluation form of acceptability of technology and procedures</td>
</tr>
</tbody>
</table>

[^a]: MIS: memory impairment screen.
[^b]: FOF: fear of falling.
[^c]: FES-I: Fall Efficacy Scale International.
[^d]: ICC: intraclass correlation coefficient.
[^e]: SEM: standard error of measurement.
InBody S10 measures impedance at 6 different frequencies (1, 5, 50, 250, 500, and 1000 kHz) for each body segment (right arm, left arm, trunk, right leg, and left leg). The spectrum of electrical frequencies is used in the manufacturer’s equation to estimate fat mass, muscle mass, and body water levels. There are no risks, no dunking, no pinching, and no discomfort associated with the use of BIA. The test duration is 1 to 2 minutes. The reliability of the BIA test-retest was high, with an ICC of 0.89 [45].

PA will be measured using activity monitoring devices. All older adults will wear ActiGraph GT9X Link wireless activity monitor (ActiGraph LLC), a triaxial accelerometer, on the nondominant wrist for 7 consecutive days. The GT9X Link has a sample rate of 30 to 100 Hz, a dynamic range of ±8G, 14-day battery life (rechargeable), 180 days/4 GB data storage, and water resistance. Data are collected at 1-minute intervals. A sensor determines whether the device is on or off the wrist. The GT9X Link provides objective 24-hour PA measures, including steps, energy expenditure, activity intensity, and participants’ posture. Accelerometry is a reliable method for assessing free-living PA (ICC=0.98) [46] and has been validated against direct observation, energy expenditure, and sedentary behavior [47,48]. ActiGraph accelerometers have been used for data collection in the National Health and Nutrition Examination surveys and are the most commonly used devices in research studies [28]. The device display screen can be disabled so that the device does not display the participant’s activity but only shows the date and time.

**Recruitment Methods**

**Identification**

We have created successful, continuing relationships with older adults and staff at the planned sites and demonstrated our ability to recruit a diverse sample [49]. Staff from our community partners and clinical sites will introduce the researchers to older adults to perform the initial screening and determine study eligibility using a checklist.

**Recruitment**

Flyers will be posted at our sites, and we will participate in community activities for face-to-face recruitment. Older adults will also be recruited from their homes by local newsletters and word of mouth. The participants will complete an informed consent process before data collection. The research team will maintain a log that tracks screening, eligibility, contact, and recruitment.

**Study Procedures and Data Collection**

**Training Staff**

All study personnel, including research assistants, will complete the required Collaborative Institutional Training Initiative training for the protection of human subjects. Training will be conducted by the authors (LT, JS, and JP), focusing on research with older adults in low-income settings, communication skills, setting up home visits, study protocols, completing questionnaires, performing tests, and contacting health care providers for older adults identified with a high level of depressive symptoms. We have research assistants who are bilingual and interested in research activities with Hispanic individuals. The authors (LT and JS) will train research assistants to perform balance and body composition tests. The third author (JP) will train research assistants to use accelerometer-based PA devices.

**Data Collection**

Following consent, participants will be contacted, and our team will meet with them to complete the questionnaires (eg, sociodemographic, FOF, or Patient Health Questionnaire-9) and the BBS balance performance test. Older adults will take off their shoes and stand as still as possible on the balance plate with their hands on their hips and eyes closed for 2 to 3 minutes. The participants will be stratified by the FRA matrix [12]. We will continually evaluate across the following 4 groups (rational, incongruent, irrational, and congruent) to maintain equal cell distribution (30 older adults per group):

- **Rational FRA**: low FOF (Short Fall Efficacy Scale International [FES-I] score ≤10) and aligned with normal balance (BBS score=0-30).
- **Incongruent FRA**: low FOF (Short FES-I score ≤10) despite poor balance (BBS score=31-100).
- **Irrational FRA**: high FOF (Short FES-I score >10) and normal balance (BBS score=0-30).
- **Congruent FRA**: high FOF (Short FES-I score >10) and poor balance (BBS score=31-100) [12].

In addition, participants will perform hand grip strength and dynamic balance performance tests. Hand grip strength will be measured in kilograms as the maximal isometric force achieved on a hand grip dynamometer. This test will be administered to participants sitting in a chair with their feet flat on the floor and their elbow bent at 90°. The dynamometer will be placed in the hand and adjusted so that the palm side of the grip will be at the palm and the front end will be lined up between the joints of the medial and distal phalanges. The grip size will be adjusted so that the second metacarpals are flat with a 90° bend at the knuckles. Participants will be asked to squeeze the strength gauge as hard as possible for 3 to 5 seconds. A total of 3 trials on each hand will be performed with a 30-second rest given between trials. Participants will complete the sit-to-stand test by standing up from a chair as much as possible within 30 seconds [50].

Participants will then be asked to wear the ActiGraph GT9X Link wireless activity monitor on the nondominant wrist for 7 consecutive days. Each participant will be asked to continue their normal activities while wearing the device. The participants will also be instructed and prepared for the body composition test during the second visit. They will be instructed to avoid exercising for 6 to 12 hours, eating for 3 to 4 hours, drinking alcohol or coffee for 24 hours, using shower or sauna, and using lotion or ointment on their hands or feet before the test. Written instructions will be provided to the participants for the use of ActiGraph and body composition tests. The questionnaires and instructions will be provided in English or Spanish by bilingual research assistants who will be trained and randomly monitored to ensure quality and consistency in administering the questionnaires and tests. The participants will be provided with
a phone number of the study team if they have questions about using the activity monitoring device.

After 7 days of wearing ActiGraph, the research assistants will collect the ActiGraph GT9X Link device. Body composition testing will occur late in the morning, and all participants will be asked to empty their bladder, remove socks, shoes, and metal objects (eg, watches and jewelry) before testing. To ensure that the body is close to a resting state as possible, participants will lie down for 15 minutes before testing in the supine position. If participants select to be tested in a seated or standing position, they will be asked to sit or stand for 10 to 15 minutes before testing (based on participants’ ability and preference). Finally, participants will be asked (using our evaluation form) about their reactions to the questions and technology (ie, what they thought about the questionnaires and technology) and asked to describe any concerns or problems in wearing the device (eg, uncomfortable to wear during walking). We will provide the BBS and BIA test results, and participants will receive Can $37.43 (US $30) store gift cards upon completion of the study.

**Participant Retention**

The following strategies will be used to maximize the participants’ retention: (1) obtaining backup contact information (eg, cell phone number or email address); (2) obtaining name and contact information for a person familiar with the participant (eg, family member or neighbor); (3) making confirmation phone calls 2 days before the second visit; (4) assigning the same data collector whenever possible; and (5) making scripted check-in calls, as needed.

**Data Management and Analysis**

**Data Management and Integrity**

The research assistants will enter deidentified questionnaire data into the Qualtrics Research Suite, a web-based survey tool that allows users to build and analyze web-based surveys and export data in multiple formats, including SPSS. Confidentiality will be maintained using coded data materials. Study data will be maintained on a dedicated server with password-protected access for authorized study staff and encrypted files. The research assistants will assist with data cleaning, preliminary analysis, and data preparation under the supervision of the first and fourth authors (LT and XY). We anticipate no more than 3% missing values on any one item, as data will be collected in-person on-site at 2 timepoints. The BIA device is supported by a custom software application, which connects to a virtual web-based data platform via a PC. The second author (JS) will conduct these analyses in collaboration with the first and fourth authors (LT and XY).

The ActiLife, the ActiGraph’s premier Actigraphy data analysis software platform, will be used to prepare the ActiGraph GT9X Link activity monitors for PA data collection and to download, score, and securely manage the collected data. ActiLife contains a wide selection of algorithms, including steps; energy expenditure; activity bouts such as standing, sitting, or lying down; sedentary analysis; and whether the ActiGraph device has been removed. A bout of sedentary behavior will be defined as consecutive minutes during which the ActiGraph registers less than 100 counts per minute. A break in sedentary behavior will be defined as at least 1 minute in which the counts registered are at least 100, following a sedentary bout. Activity counts and step data will be used to derive a series of PA and sedentary behavior indicator variables. We will assess the feasibility of wearing the devices by assessing the frequency of older adults who have at least 4 days of at least 10 hours of wear per day (detected using device technology), which is the standard convention [51]. The third author (JP) will conduct these analyses in collaboration with the first and fourth authors (LT and XY).

**Data Analysis**

All analyses will be performed in SPSS version 25 or SAS version 9.4, with sufficient annotation for reproducibility. Descriptive and exploratory analyses will be performed first to investigate the distributional assumptions. Although no differences are hypothesized, descriptive subgroup analyses by sex will be conducted.

**Aim 1: Examine the Feasibility of Recruitment and Acceptability of Technologies and Procedures for Use Among Older Adults in Low-Income Settings**

We will assess the ability to recruit the sample by calculating the proportions of older adults with low income (1) who are recruited out of the total screened and (2) who completed all study procedures. We will track the number of days and the time spent to recruit the sample. The results will inform the planning for future larger studies. The acceptability of the technologies and procedures will be examined based on an evaluation form (eg, what they thought about the questionnaires and technology) and their recommendations. We will identify the code, categorize participants’ responses, and determine the frequency for each category. Demographic data, including essential characteristics of older adults, will be obtained from the study participants. We anticipate no more than 3% missing values for any one item, as data will be collected in-person on-site.

**Aim 2: Examine the Associations Among FRA, Body Composition, and PA**

The hypotheses of the second aim are as follows: (1) rational FRA is associated with higher levels of PA and skeletal muscle mass and with lower levels of percentage of body fat and BMI, (2) incongruent FRA is associated with higher levels of PA and skeletal muscle mass and with lower levels of percentage of body fat and BMI, and (3) irrational and congruent FRAs are associated with lower levels of PA and skeletal muscle mass and with higher levels of percentage of body fat and BMI.

General summary statistics will be presented for continuous data, and percentages will be presented for categorical data. The participants’ demographic data will be summarized. Analyses will be performed without adjusting for any covariates, followed by analyses adjusted for age, comorbidities, depressive symptoms, and fall history. The number of comorbidities and the number of falls in the past year will be selected as covariates in the selected model because of their association with FRA, body composition, and PA. Data from InBodys10 (eg, BMI) and ActiGraph (eg, steps per day) could be modeled as continuous.
Regression models or one-way ANOVA (analysis of variance) will be used to examine the differences in continuous variables (eg, BMI and steps per day) across the 4 groups (rational, incongruent, irrational, and congruent FRAs), and categorical variables (eg, comorbidities) will be tested using chi-square tests. The rational FRA group will set as a base group, and the other 3 groups (incongruent, irrational, and congruent FRAs) will be compared with this base group. Whenever feasible, models will be adjusted for age, sex, day order, wear time (min per day), and fall history. The following variables will be added one by one to evaluate the role of mobility problems (yes or no), depression (score or category), fall risk (high or low), perceived general health (score or category), medication use (yes or no), and comorbidities (yes or no). A final model will include all significant potential variables to evaluate the associations among FRA, body composition, and PA, whenever feasible.

Sample Size and Power Analysis

There is no historical precedent, and there is limited information to inform the sample size for this study; therefore, the sample size estimate was based on the variable of the BBS score in our pilot study. The mean BBS scores in our pilot study were 26.44 (SD 7.91) in the high PA group and 30.62 (SD 7.78) in the low PA group (PA by self-report), and assuming a two-tail $\alpha$ of .05 and a power of 0.8, the estimated total sample size of 120 would allow sufficient detection of a mean difference of about 5 hypothetical units of BBS score. Furthermore, to compare the effect sizes of the different body compositions and PA measures, we performed posthoc analyses to compare the mean values between the 2 groups of PA, adjusted for the age group.

Potential Challenges and Solutions

Recruitment and Retention

Although recruitment of diverse older adults to research studies is often challenging, our tailored strategies will address this issue by ensuring that our diverse research assistants are well trained and by employing a straightforward data collection process.

Respondent Burden

We will use questionnaires in either the English or Spanish version, based on the preferences of the older adults. Research assistants will read the questions and help older adults complete if they cannot do so.

Risk of Falls During Tests

We will follow the test protocol, which was used successfully in our pilot study; we had no falls when using the BBS.

Ethical Approval and Consent to Participate

Ethical approval was obtained from the institutional review board of the University of Central Florida (protocol number 2189; September 10, 2020). All participants will be informed before their participation.

Results

The project funding was received on September 30, 2020, and was approved by the Institutional Review Board (September 10, 2020). As of March 2021, we have enrolled 11 participants and expect the results to be published in Summer 2022.

Discussion

Accurate FRA is essential for implementing PA programs, especially for older adults with low income. Systematic reviews found that most studies have not used objective measures of fall risk and PA [5,52]. Most studies have used self-reported measures of fall risk and PA among older adults [53-57]. Older adults who overestimate their fall risk are less likely to participate in PA [58,59]. Using AHT may eliminate recall bias associated with subjective measures [60,61]. Systematic reviews also indicate the importance of technology-based assessments for PA and to prevent falls [62,63]. This study will provide data for developing technology-based fall risk assessments to improve participation in PA, thus enhancing healthy longevity among older adults in low-income settings.

Acknowledgments

This project received financial support from the National Institute on Aging, National Institutes of Health, grant number 1R03AG06799-0, with the projected title, “Technology-based fall risk assessments for older adults in low-income settings.” This project is also supported by the Learning Institute for Elders at the University of Central Florida. The funding body is not involved in the study design, data collection, data analysis, or writing of the manuscript.

Authors’ Contributions

LT, JS, JP, and XY contributed to the study conceptualization, design, and data analysis. LT wrote the original draft, and JS, JP, and XY contributed to the substantial revision of the original draft. All authors have agreed to the final version of the paper.

Conflicts of Interest

None declared.

References


43. Goble DJ. BTrackSTM Balance Test is a valid predictor of older adults falling. 2018. URL: https://balancetrackingsystems.com/assess-balance/ [accessed 2021-03-12]


Abbreviations

AHT: assistive health technology
BBP: BTrackS balance plate
BBS: BTrackS balance system
BIA: bioelectrical impedance analysis
FES-I: Fall Efficacy Scale International
FOF: fear of falling
FRA: fall risk appraisal
ICC: intraclass correlation coefficient
PA: physical activity
Effects of Participating in a Research Project During the COVID-19 Pandemic on Medical Students’ Educational Routines and Mental Health: Protocol for a Web-Based Survey Study

Débora Cerqueira Calderaro1*, MD, PhD; Barbara Stadler Kahlow2*, MSc, MD; Gabriela Araújo Munhoz3*, MD, PhD; Samuel Elias Basulto Dias4*, MS; João Vitor Ziroldo Lopes5*, MS; Aline Rizzo Borges6*, MS; Henrique De Ataíde Mariz7*, MD, PhD; Kirla Wagner Poti Gomes8*, MD, PhD; Lilian David De Azevedo Valadares7*, MSc, MD; Nafice Costa Araújo9*, MSc, MD; Sandra Lucia Euzébio Ribeiro4*, MD, PhD; Adriana Maria Kakehashi1*, MD, PhD; Ana Paula Monteiro Gomides Reis10*, MD, PhD; Cláudia Marques7*, MD, PhD; Edgard Torres Reis-Neto11*, MD, PhD; Eduardo Dos Santos Paiva12*, MD, PhD; Gecilmara Salviato Pileggi13*, MD, PhD; Gilda Aparecida Ferreira14*, MD, PhD; José Roberto Provenza14*, MD, PhD; Lícia Maria Henrique Mota10*, MD, PhD; Ricardo Machado Xavier15*, MD, PhD; Maycoln Leôni Martins Teodoro15*, MD, PhD; Marcelo De Medeiros Pinheiro11*, MD, PhD; The Voluntary Brazilian Society Of Rheumatology Task Force Against COVID-1916*

1 Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
2 Rheumatology Department, Hospital Universitário Evangélico Mackenzie, Paraná, Brazil
3 Rheumatology Department, Santa Casa de Misericórdia de São Paulo, São Paulo, Brazil
4 Rheumatology Department, Universidade Federal do Amazonas, Manaus, Brazil
5 Medical School, Universidade de São Paulo, São Paulo, Brazil
6 Medical School, Centro Universitário de Brasília, Brasília, Brazil
7 Rheumatology Department, Universidade Federal de Pernambuco, Recife, Brazil
8 Rheumatology Department, Universidade Federal de Fortaleza, Fortaleza, Brazil
9 Rheumatology Department, Hospital do Servidor Público Estadual, Instituto de Assistência Médica ao Servidor Público Estadual, São Paulo, Brazil
10 Rheumatology Department, Universidade de Brasília, Brasília, Brazil
11 Rheumatology Department, Universidade Federal de São Paulo, São Paulo, Brazil
12 Rheumatology Department, Universidade Federal do Paraná, Curitiba, Brazil
13 Rheumatology Department, Faculdade de Ciências da Saúde de Barretos, Barretos, Brazil
14 Rheumatology Department, Pontifícia Universidade Católica de Campinas, Campinas, Brazil
15 Rheumatology Department, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
16 see Acknowledgments

* all authors contributed equally

Corresponding Author: Débora Cerqueira Calderaro, MD, PhD
Universidade Federal de Minas Gerais
190 Professor Alfredo Balena Avenue
Belo Horizonte, 30130-100
Brazil
Phone: 55 3134099757
Email: dccalderaro@gmail.com

Abstract

Background: The COVID-19 pandemic has resulted in social isolation, which has a potential negative impact on the educational routines (eg, the suspension of face-to-face appointments) and mental health of medical students. The Mario Pinotti II (MPII) study is a 24-week observational study that conducted scheduled telephone calls every 2 weeks to verify the occurrence of COVID-19 in patients with rheumatic diseases on chronic hydroxychloroquine therapy (from March 29, 2020, to September 30, 2020).
2020). The effects of voluntarily participating in a research project (ie, one that involves interactions via telephone contact with patients, professors, rheumatologists, and colleagues) on the daily lives and mental health of medical students requires evaluation.

**Objective:** As medical students are professionals in training and have a high level of responsibility in terms of handling the emotional and physical aspects of several diseases, this study aims to evaluate the impacts of the COVID-19 pandemic and participation in the MPII study on the educational routines and mental health of medical students.

**Methods:** A web-based survey was carried out to perform a cross-sectional comparative assessment of medical students who participated in the MPII study and their colleagues who were not involved in the MPII study. Participants from both groups were matched based on sex, age, and medical school. The web questionnaire was developed by a panel composed of graduate medical students, rheumatologists, medical school professors, and a psychology professor. The questionnaire included details on demographic and life habits data and evaluated participants' impressions of the MPII study and the impact of the COVID-19 pandemic on their educational routines and medical training. In addition, depression, anxiety, and stress were evaluated using the Brazilian version of the Depression, Anxiety, and Stress Scale (DASS)-21, and currently, the DASS-21 scores are grouped as those that indicate a low, moderate, or high risk of mental distress. This project was approved by the Federal University of São Paulo Ethics Committee (CAAE: 34034620.0.0000.5505).

**Results:** Data were collected from both medical student groups from July 20 to August 31, 2020. Data extraction was completed in September 2020. The data analysis is ongoing. We expect the results to be published in the first semester of 2021.

**Conclusions:** This study will provide insight into the effects of participating in a research project on depression, anxiety, and stress, which will be determined by applying the DASS-21 to a large sample of Brazilian undergraduate medical students. We will also evaluate the impact of the COVID-19 pandemic on medical students’ educational routines and medical training.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/24617

*(JMIR Res Protoc 2021;10(4):e24617) doi:10.2196/24617*

**KEYWORDS**

SARS-CoV-2; COVID-19; medical education; observational; cross-sectional; case-control study; voluntary; mental health; rheumatic disease; medical student; protocol; survey

**Introduction**

In December 2019, a respiratory disease (COVID-19) that is caused by the novel SARS-CoV-2 was identified in Wuhan City, Hubei Province, China. Several weeks later, the World Health Organization declared COVID-19 as an international public health emergency and pandemic [1].

Given the high levels of community transmission of SARS-CoV-2, several approaches have been recommended to mitigate global viral spread; social distancing, quarantine, intermittent hand hygiene, and universal masking are especially important approaches. However, these government recommendations vary according to each country and have resulted in home isolation, fear, uncertainty, anxiety, depression, high alcohol intake, domestic violence, education impairment, and severe economic burdens [2-4].

With regard to medical education, the majority of medical classes and face-to-face practices were suspended for a long period while medical schools prepared for remote education. Furthermore, several students have been invited to work on the front line of the pandemic (ie, treating patients with COVID-19), which allowed them to graduate earlier [5,6]. These measures, which are associated with remote (not in-person) training, have hampered medical training for working with patients and other relevant aspects, including hospital and outpatient clinics and regulatory processes [7]. Several researchers have highlighted a gradual increase in anxiety levels among medical students during the COVID-19 pandemic, suggesting that the pandemic may be impairing several aspects of social relationships, technical performance, and mental health [7,8].

It is worth emphasizing that medical training is often exhausting due to specific technical requirements and a large number of stressful factors [9], including full-time dedication, personal life–related effort and sacrifice, close contact with severe diseases and death, and physical and emotional distress [10]. During the COVID-19 pandemic, extensive and intense workloads, difficulties in reconciling personal life with studies, competitiveness, sleep deprivation, fears of making mistakes and getting sick, tiredness, and decision making under pressure may result in high levels of anxiety and depression in medical students [9-13].

European studies have reported that around 30% of medical students experience some level of depression or anxiety. In Brazil, studies have suggested that 20%-50% of medical students experience mood changes [12]. In addition, depression and suicidal ideation rates are higher in medical students than in the general population, and these students generally seek less help from psychological or psychiatric professionals [5-12]. Several psychiatric illnesses and personality disturbances have been reported to be related to such behavior, including eating disorders, the denial of reality, alcoholism, the abuse of illicit drugs, a lack of commitment, obsessive-compulsive disorder, anxiety, depression, and increased suicide rates [7]. Thus, medical students are susceptible to experiencing inadequate or nonadaptive responses to emotional distress [7,13].

Since January 2020, the World Health Organization has been warning the public that the COVID-19 pandemic is generating...
stress in the general population [1], especially stress related to uncertainties about the course and prognosis of the disease; fear; a lack of resources for diagnosis and treatment; a shortage of food, medication, and adequate supplies of personal protective equipment for health care professionals; feelings of missing the freedom of travel; and conflicting information delivered by governmental authorities or social media [14-16]. It has also been reported that the incidence of several psychological disorders increased during previous pandemics, including the SARS (Severe Acute Respiratory Syndrome and MERS (Middle Eastern Respiratory Syndrome) pandemics. Such disorders mainly included anxiety and depression [8].

During the COVID-19 pandemic in China, a study that evaluated 217 medical students reported that depression and anxiety occurred in 35% and 22% of students, respectively [8]. Furthermore, researchers at medical schools in Wisconsin believe that students are great allies of doctors [7].

The Mario Pinotti II (MPII) study is a noninterventional, observational, multicenter, parallel-group cohort study that included adult volunteers (aged ≥18 years) with a previously known diagnosis of rheumatic disease. These participants were on hydroxychloroquine for at least 30 days prior to baseline. The MPII is a 24-week prospective study that included more than 10,000 individuals from 20 centers in Brazil. A total of 6 sequential telephone calls were scheduled during the community transmission of SARS-CoV-2. These calls were performed by 395 volunteer medical students [17].

Given the close social interaction between the patients and controls, which was facilitated through periodic telephone contact, as well as the social interactions among principal investigators, study coordinators, and professors, our main hypothesis was that medical students who participated in the MPII study would experience less emotional distress than their colleagues who did not participate in the MPII study. [12-15].

Our objectives were to evaluate the impact of participating in the MPII study on the mental health (evaluated using the Depression, Anxiety, and Stress Scale [DASS]-21) [17,19], professional improvement, and commitment perceptions of medical students during the COVID-19 pandemic. We also aimed to identify potential impairments in the educational routines of students’ medical schools and to report on COVID-19 diagnoses among this population.

Methods

Study Design

We will conduct a comparative, cross-sectional, observational, case-control study that used a voluntary web-based survey. The survey was conducted according to the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) statement [20].

Study Population

Medical students who were involved in the MPII study (volunteer group) and their colleagues who were not involved in the MPII study (control group) [17] were recruited.

Sample Size

Convenience sampling was conducted based on voluntary involvement, as per the investigators in the MPII Study. Of the 20 MPII centers, 14 (70%) participated in this study. All students who successfully answered the web questionnaire during the data collection period were included in this study.

Inclusion Criteria

The volunteer group consisted of medical students who participated as volunteers in the MPII study and were aged ≥18 years. The control group consisted of medical students who did not participate in the MPII study. For each participant in the volunteer group, at least two other students were enrolled as controls. This was done to ensure that the number of participants was sufficient for stratified analysis, as we expected that the prevalence of mental health issues among medical students would be high [21].

An electronic informed consent form was provided to both groups, and participants were required to sign it before they were granted access to the survey questionnaire.

Exclusion Criteria

Participants were excluded if they refused to provide informed consent or withdrew their consent.

Survey Questionnaire

This study was approved by the Federal University of São Paulo Ethics Committee (Certificado de Apresentação de Apreciaçãoética [CAAE]: 34034620.0.0000.5505) on July 13, 2020.

The complete survey questionnaire form can be found in Multimedia Appendix 1. It includes 69 questions. The time required for completing the survey was 20 minutes. The questionnaire was developed and provided to participants in Portuguese (their mother language), and it was only translated in order to be published. The translation was verified by an experienced translator in Brazil.

A panel of undergraduate medical students, rheumatologists, and medical school professors who were involved in the MPII study, as well as a psychology professor who had experience in conducting web-based surveys to evaluate the mental health of medical students and health care professionals, were responsible for developing the web-based survey questionnaire. The students in the panel that developed the survey (n=3) tested the questionnaire. Afterward, it was distributed to the other participants of this study.

The volunteer group received an invitation video that provided explanations about the survey and the link to access the web questionnaire. Subsequently, the volunteers were requested to send an invitation link to colleagues who were not participating in the MPII study. These colleagues were added to the control group. No identification data were requested.

An informed consent form was integrated into the web questionnaire. Participants were required to provide their consent to participate in the survey (electronic informed consent) before accessing the questionnaire of the study or providing any information.
Participants can provide an email address if they want to receive their mental health evaluation results. Participants or researchers who require a psychological or psychiatric evaluation will be provided (through email contact) with guidance on accessing local facilities that offer these services. All participants were informed that providing an email address is voluntary and can be done before answering the survey questionnaire.

Demographic and epidemiologic data and details about comorbidities, life habits (smoking, alcohol intake, illicit drug use, and physical activity), and concomitant medications were recorded. In addition, specific aspects related to medical schools, such as the type of school (public or private), costs, teaching activities, and feelings about medical training during the pandemic, were addressed. With regard to participants in the volunteer group, their impressions of the procedures in the MPII study and the study’s impact on their daily routines were evaluated.

The Brazilian version of the DASS-21 was used to evaluate mental health [18,19]. The DASS-21 is a set of three self-report scales with seven items each. The items were designed to measure emotional status. The depression domain assesses dysphoria, hopelessness, the devaluation of life, self-deprecation, a lack of interest or commitment, anhedonia, and inertia. The anxiety domain evaluates autonomic symptoms, skeletal muscle effects, coping, and experiences. The stress domain is sensitive to levels of chronic nonspecific arousal, and it assesses difficulties in relaxing; nervous arousal; and the state of being easily upset/agitated, irritable/overreactive, and impatient. Scores for depression, anxiety, and stress are calculated by summing the scores of the relevant items. The DASS-21 is based on a dimensional concept of psychological disorders instead of a categorical concept of psychological disorders, and it was developed by accounting for the differences among the depression, anxiety, and stress experienced by subjects. Therefore, the scale does not have any direct implications for diagnosis.

DASS-21 scores were grouped as those that indicate a low, moderate, or high risk or mental distress. These will be stratified according to gender and the SD from the study population’s mean scores (low: lower than mean+1 SD; moderate: ranges from mean+1 SD to mean+2 SDs; high: greater than mean+2 SDs).

Recommended cutoff scores for conventional severity labels of depression, anxiety, and stress (normal, mild, moderate, severe, and extremely severe) were multiplied by 2, as shown in Table 1 [19].

Table 1. Cutoff scores for depression, anxiety, and stress according to the Depression, Anxiety, and Stress Scale (DASS)-21.

<table>
<thead>
<tr>
<th>Severity label</th>
<th>Doubled DASS-21 domain scores [19]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Depression</td>
</tr>
<tr>
<td>Normal</td>
<td>0-9</td>
</tr>
<tr>
<td>Mild</td>
<td>10-13</td>
</tr>
<tr>
<td>Moderate</td>
<td>14-20</td>
</tr>
<tr>
<td>Severe</td>
<td>21-27</td>
</tr>
<tr>
<td>Extremely severe</td>
<td>≥28</td>
</tr>
</tbody>
</table>

Data Collection

Data collection was performed by using a web questionnaire that was generated on the Google Forms platform. The questionnaire was disclosed to the research subjects via email and WhatsApp (WhatsApp LLC). The data collection period was from July 20 to August 31, 2020.

Statistical Analysis

Descriptive analysis will be performed using absolute and relative frequencies for categorical variables and quantitative measures (means, quartiles, minimums, maximums, and SDs) for numerical variables. The normality of numerical variables will be evaluated using the Kolmogorov-Smirnov test. Numerical variables with normal distributions will be described as mean (SD), and nonnormal numerical variables will be described as median (IQR) or range (minimum-maximum).

The Chi-square association test with adjusted standardized residuals will be used to assess the association between categorical variables; the Fisher exact test will be used for small samples. The linear associations between two numerical variables will be evaluated using the Pearson or Spearman correlation method.

Adjusted multiple linear regression models will be used to assess the simultaneous effects of sex, age, comorbidities, concomitant medications, and other confounding variables based on group type and predefined outcomes (anxiety, depression, and stress scores from the DASS-21). For dichotomous dependent variables, a logistic regression model will be used.

SPSS, version 20 (IBM Corporation) will be used for all analyses. A P value of <.05 will be considered significant.

Results

Data were collected from both medical student groups from July 20 to August 31, 2020. Data extraction was completed on September 2020. The data analysis is ongoing. We expect the results to be published in the first semester of 2021.
Discussion

This study has an unprecedented design, as it includes a very large sample of volunteer medical students from 14 Brazilian tertiary rheumatology centers. These students are currently monitoring the outcomes of 9589 patients with rheumatic diseases on hydroxychloroquine and assessing patients’ susceptibility to SARS-CoV-2 infection.

The main objective of this study is to evaluate the impact of participating in a research project during the COVID-19 outbreak on the mental health and learning behaviors of medical students. These variables were measured by using a structured web questionnaire about students’ volunteer participation in the MPII study and their ability to work with patients and professors in a real-life scenario.

This study has several innovative aspects, such as (1) the evaluation of depression, anxiety, and stress by applying the DASS-21 to a large sample of medical students and a control group; (2) medical students’ impressions regarding the handling of the uncertainty and doubts of patients with rheumatic diseases, including the fear of illness, fear of dying, and shortage of medication during the outbreak; (3) the measurement of the impact of the COVID-19 pandemic on students’ educational routines and medical training; and (4) the fact that the web questionnaire was developed by a panel composed of graduate medical students, rheumatologists, medical school professors, and a psychology professor.

Acknowledgments

We would like to acknowledge Faculdade de Medicina de Barretos, Education & Research Institute Cancer Hospital Barretos, Barretos-SP, Brazil; Hospital das Clínicas, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil; Hospital Universitário de Brasília da Universidade de Brasília, Empresa Brasileira de Serviços Hospitalares, Brasília-DF, Brazil; Hospital São Paulo da Universidade Federal de São Paulo, Escola Paulista de Medicina, São Paulo-SP, Brazil; Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brazil; Hospital do Servidor Público Estadual, Instituto de Assistência Médica ao Servidor Público Estadual, São Paulo-SP, Brazil; Hospital Universitário Lauro Wanderley, Universidade Federal da Paraíba, João Pessoa-PB, Brazil; Hospital das Clínicas da Universidade Federal de Pernambuco, Recife-PE, Brazil; Hospital Universitário da Universidade Federal de Juiz de Fora, Juiz de Fora-MG, Brazil; Hospital Geral de Fortaleza, Fortaleza-CE, Brazil; Hospital Universitário Pedro Ernesto, Universidade do Estado do Rio de Janeiro, Rio de Janeiro-RJ, Brazil; Irmandade da Santa Casa de Misericórdia de São Paulo, São Paulo-SP, Brazil; Universidade Federal do Paraná, Curitiba-PR, Brazil; Hospital Universitário Getúlio Vargas Universidade Federal do Amazonas, Manaus-AM, Brazil; Hospital Universitário Evangélico Mackenzie, Curitiba-PR, Brazil; and Hospital Universitário Caxias de Mares, Universidade Federal do Espírito Santo, Vitória-ES, Brazil. This paper was written on behalf of The Voluntary Brazilian Society of Rheumatology Task Force Against COVID-19.

The members of The Voluntary Brazilian Society of Rheumatology Task Force Against COVID-19, Mirthelen Abreu, Cleandros Albuquerque, Ana Bacchiega, Blanca Bica, Danielle Brito, Ângela Duarte, Paula Fernandes, Ana Guedes, Evandro Klumb, Helena Pereira, Emília Sato, Thelma Skare, Viviane Souza, Valéria Valim and Elaine Bezerra collaborated on this work, contributing to its conception, spreading the in-survey questionnaire, and data collection.

Authors’ Contributions

Authors DCC, BSK, GAM, SEBD, JVZL, ARB, HDAM, KWPG, LDAV, NCA, SLER, AMK, APMGR, CM, ETR-N, EDSP, GSP, GAF, JRP, LMHM, RMX, MLMT, and MDMP have contributed to the study conception and protocol design, drafting and critically reviewing this manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1 Survey instrument. [PDF File (Adobe PDF File), 67 KB, resprot_v10i4e24617_app1.pdf]

References


15. Estrela YDCA, Rezende ACC, Guedes AF, Pereira CDO, de Sousa MNA. Estresse e correlatos com características de saúde e sociodemográficas de estudantes de medicina. Estress y características sociodemográficas y de salud en estudiantes de medicina. CES Medicina 2018 Dec;32(3):215-225 [FREE Full text] [doi: 10.21615/cesmedicina.32.3.3]


Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
DASS: Depression, Anxiety, and Stress Scale
MPII: Mario Pinotti II
MERS: Middle Eastern Respiratory Syndrome
SARS: Severe Acute Respiratory Syndrome

https://www.researchprotocols.org/2021/4/e24617
JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e24617 | p.255
(page number not for citation purposes)
Edited by G Eysenbach; submitted 28.09.20; peer-reviewed by E Oghogho, P Kanzow; comments to author 08.03.21; revised version received 11.03.21; accepted 17.03.21; published 09.04.21.

Please cite as:
JMIR Res Protoc 2021;10(4):e24617
URL: https://www.researchprotocols.org/2021/4/e24617
doi:10.2196/24617
PMID:33735094

©Débora Cerqueira Calderaro, Barbara Stadler Kahlow, Gabriela Araújo Munhoz, Samuel Elias Basualto Dias, João Vítor Ziroldo Lopes, Aline Rizzo Borges, Henrique De Ataíde Mariz, Kirla Wagner Poti Gomes, Lilian David De Azevedo Valadares, Nafice Costa Araújo, Sandra Lucia Euzébio Ribeiro, Adriana Maria Kakehasi, Ana Paula Monteiro Gomides Reis, Cláudia Marques, Edgard Torres Reis-Neto, Eduardo Dos Santos Paiva, Gecilmara Salvato Pileggi, Gilda Aparecida Ferreira, José Roberto Provenza, Licia Maria Henrique Mota, Ricardo Machado Xavier, Maycoln Leôni Martins Teodoro, Marcelo De Medeiros Pinheiro, The Voluntary Brazilian Society Of Rheumatology Task Force Against COVID-19. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 09.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

Generating Intervention Concepts for Reducing Adolescent Relationship Abuse Inequities Among Sexual and Gender Minority Youth: Protocol for a Web-Based, Longitudinal, Human-Centered Design Study

Robert W S Coulter1,2,3, PhD, MPH; Shannon Mitchell1, BS; Kelly Prangley1, BS; Seth Smallwood1; Leyna Bonanno1, MS; Elizabeth N Foster1, MLiS; Abby Wilson4, JD, LLM; Elizabeth Miller2,3, MD, PhD; Carla D Chugani2,3, PhD, LPC

1Department of Behavioral and Community Health Sciences, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States
2Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States
3Division of Adolescent and Young Adult Medicine, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, United States
4LUMA Institute, Pittsburgh, PA, United States

Corresponding Author:
Robert W S Coulter, PhD, MPH
Department of Behavioral and Community Health Sciences
Graduate School of Public Health
University of Pittsburgh
130 DeSoto St
Pittsburgh, PA, 15261
United States
Phone: 1 4126240647
Email: robert.ws.coulter@pitt.edu

Abstract

Background: Sexual and gender minority youth (SGMY; eg, lesbian, gay, bisexual, and transgender youth) are at greater risk than their cisgender heterosexual peers for adolescent relationship abuse (ARA; physical, sexual, or psychological abuse in a romantic relationship). However, there is a dearth of efficacious interventions for reducing ARA among SGMY. To address this intervention gap, we designed a novel web-based methodology leveraging the field of human-centered design to generate multiple ARA intervention concepts with SGMY.

Objective: This paper aims to describe study procedures for a pilot study to rigorously test the feasibility, acceptability, and appropriateness of using web-based human-centered design methods with SGMY to create novel, stakeholder-driven ARA intervention concepts.

Methods: We are conducting a longitudinal, web-based human-centered design study with 45-60 SGMY (aged between 14 and 18 years) recruited via social media from across the United States. Using MURAL (a collaborative, visual web-based workspace) and Zoom (a videoconferencing platform), the SGMY will participate in four group-based sessions (1.5 hours each). In session 1, the SGMY will use rose-thorn-bud to individually document their ideas about healthy and unhealthy relationship characteristics and then use affinity clustering as a group to categorize their self-reported ideas based on similarities and differences. In session 2, the SGMY will use rose-thorn-bud to individually critique a universal evidence-based intervention to reduce ARA and affinity clustering to aggregate their ideas as a group. In session 3, the SGMY will use a creative matrix to generate intervention ideas for reducing ARA among them and force-rank the intervention ideas based on their potential ease of implementation and potential impact using an importance-difficulty matrix. In session 4, the SGMY will generate and refine intervention concepts (from session 3 ideations) to reduce ARA using round robin (for rapid iteration) and concept poster (for fleshing out ideas more fully). We will use content analyses to document the intervention concepts. In a follow-up survey, the SGMY will complete validated measures about the feasibility, acceptability, and appropriateness of the web-based human-centered design methods (a priori benchmarks for success: means >3.75 on each 5-point scale).

Results: This study was funded in February 2020. Data collection began in August 2020 and will be completed by April 2021.
Conclusions: Through rigorous testing of the feasibility of our web-based human-centered design methodology, our study may help demonstrate the use of human-centered design methods to engage harder-to-reach stakeholders and actively involve them in the co-creation of relevant interventions. Successful completion of this project also has the potential to catalyze intervention research to address ARA inequities for SGMY. Finally, our approach may be transferable to other populations and health topics, thereby advancing prevention science and health equity.

International Registered Report Identifier (IRRID): DERR1-10.2196/26554


KEYWORDS

sexual and gender minorities; adolescent; psychosocial intervention; internet-based intervention; intimate partner violence

Introduction

Background

Sexual and gender minority youth (SGMY; eg, lesbian, gay, bisexual, and transgender youth) are at greater risk than their cisgender heterosexual peers for experiencing adolescent relationship abuse (ARA; ie, physical, sexual, or psychological abuse in a romantic relationship) [1-4]. According to a recent nationally representative sample of high school youth in the United States, 13% of sexual minority youth and only 7% of heterosexual youth reported physical ARA in the past year [5]. Even greater disparities are present for sexual ARA, with a prevalence of 16% among sexual minority youth and only 7% among heterosexual youth [5]. Gender minority youth are also at greater risk for ARA than their cisgender peers [6]. SGMY ARA inequities are problematic because ARA is associated with many poor health outcomes later in life, such as mental health disorders and HIV [7,8]. Thus, prevention efforts for SGMY may mitigate health inequities more broadly.

Despite researchers and national agencies calling for interventions to reduce ARA among SGMY [9-11], there exist few evidence-based interventions addressing this public health inequity [9,12]. While there are several efficacious interventions for reducing ARA for the entire adolescent population [13], a 2019 systematic review revealed that there were no evidence-based ARA interventions specifically for SGMY at that time [12]. More recently, one study examined the efficacy of a universal intervention for reducing ARA among sexual minority youth; for sexual minority youth, this intervention reduced stalking victimization but not sexual violence, sexual harassment, and physical dating violence victimization [14]. More research is needed to address the lack of evidence-based interventions to prevent and reduce ARA in SGMY.

One innovative method for stimulating new stakeholder-driven intervention ideas to catalyze ARA research among SGMY is to leverage the field of human-centered design [15]. Human-centered design is a discipline focused on improving existing or developing new products, services, or experiences by involving the perspectives of the target population at every possible stage [16-27]. Human-centered design methods often incorporate multiple ways of soliciting user and other stakeholder input, including through observation and dialog, cooperative design activities, and the shared creation of meaning by collaboratively synthesizing, critiquing, and ranking self-reported data and observations. In a practical sense, this often means that individuals involved in human-centered design processes create their own artifacts, including assembling, disassembling, and reassembling qualitative data points written on Post-it notes or a digital analog.

When bringing stakeholders together to design interventions, human-centered design generally harnesses the strengths and limits the weaknesses of more traditional approaches (eg, focus groups). For example, focus groups traditionally use group-based discussions and interviews generally work with participants one-on-one. Human-centered design methods, however, often combine tasks to be completed in groups with tasks to be completed as individuals, thereby harnessing the strengths of focus groups and interviews [28,29]. Focus groups are also prone to groupthink (conformity of individuals working in groups, despite their individual differences, which can lead to inaccurate results or poor decision making [30]) that may inadvertently reinforce social hierarchies that silence certain people (eg, marginalized or shy people) [31-34]. In contrast, many human-centered design techniques require each participant to brainstorm independently and record all their ideas in written format (before sharing output with other group participants), making data collection more comprehensive and equitable [28,35]. In addition, focus groups are prone to social desirability bias because the moderator has an active role in guiding and influencing the discussion [32-34,36]. In human-centered design, facilitators have a less subjective role as they usually only lead participants through task instructions without any probing [28]. Finally, conventional methods (eg, focus groups or expert panels) for bringing stakeholders to translate research findings into intervention concepts are often challenging and time-intensive [37-41]. To overcome these barriers, human-centered design uses structured activities that are time-limited, engaging, and accessible to laypersons, including youth [16-27]. Overall, human-centered design activities can be applied as a novel form of stakeholder-engaged research to rapidly generate and iterate intervention concepts to reduce emergent public health problems.

To date, health research that uses human-centered design methods has been predominantly conducted with stakeholders in-person [42], but most in-person research activities have been impeded by the COVID-19 pandemic. Although, the pandemic has presented many barriers for safely conducting in-person research, it has simultaneously catalyzed the use of web-based technologies for interacting and collaborating. Young people have especially become accustomed to using web-based technologies (eg, Zoom, a videoconferencing platform) because
many schools have transitioned to remote learning. Therefore, youth are uniquely poised to use web-based human-centered design methods. Previous research has shown that youth can feasibly engage in web-based research as well as in-person human-centered design activities [43-47]. However, to our knowledge no study has explicitly tested whether human-centered design methods are feasible, acceptable, and appropriate for engaging with youth in a fully web-based environment. A study that rigorously pilot tests such methods, by setting a priori benchmarks, can help inform the public health field about the utility of conducting these methods on the web and further demonstrate and codify the use of web-based human-centered design as a method for stakeholder-engaged research.

Study Aims

This paper describes a protocol for conducting web-based human-centered design sessions with SGMY to create novel intervention ideas for addressing ARA. The specific aims of this study are as follows:

1. Test the feasibility, acceptability, and appropriateness of conducting web-based human-centered design sessions with SGMY (primary study aim).
2. Elucidate the beliefs of SGMY about healthy and unhealthy characteristics of intimate relationships.
3. Elicit feedback from SGMY about the School Health Center Healthy Adolescent Relationships Program (SHARP), which is a universal evidence-based intervention for reducing ARA [48,49], and about adapted SHARP materials that are tailored to SGMY [50,51].
4. Brainstorm intervention ideas for reducing ARA inequities for SGMY and force-rank the intervention ideas based on their potential ease of implementation and potential impact.
5. Generate, iterate, vote on, and refine the intervention concepts to reduce ARA inequities for SGMY.

Methods

Overview

We followed both the Standards for Reporting Qualitative Research [52] and the Strengthening the Reporting of Observational Studies in Epidemiology statement [53] to craft our study, present our methods, and report our results. We used these two guidelines because there are no formal reporting requirements for human-centered design studies.

Study Design

We are conducting a web-based, longitudinal human-centered design study to engage with 45 to 60 SGMY participants in small groups to generate intervention concepts for reducing SGMY inequities in ARA. We conducted longitudinal sessions in multiple cohorts of 8 to 15 participants each. Participants completed a web-based screener, four web-based, group-based human-centered design sessions, and a web-based follow-up survey (Figure 1). This study is primarily funded by the National Center for Advancing Translational Sciences at the National Institutes of Health (UL1TR001857).

Study Population

We aim to enroll 45 to 60 sexual and gender minority high school students recruited via social media advertisements. Eligible youth are aged 14 years to 18 years, live in the United States, identify as sexual and/or gender minorities, and have internet, video camera, audio, and microphone access to attend the virtual sessions.

Recruitment

Participants are conveniently sampled and recruited throughout the United States using web-based advertisements posted on two social media platforms, Facebook and Instagram, using an approach similar to that in our previous research [54]. This approach allows SGMY from multiple geographic locations (eg, rural and urban areas and East and West) to enroll in the study without overextending our resources. Facebook is an appropriate recruitment platform because it is highly used by adolescents, approximately 71% of teens use Facebook [55]. Similarly, Instagram is used by approximately 72% of teens in the United States, with the majority using the site daily [56]. We created multiple photo and video advertisements for these sites, including those depicting youth with diverse gender expressions, race, and ethnicities.

Screener Survey

Upon clicking the advertisements, potential participants are redirected to a brief web-based self-reported screener survey administered via Research Electronic Data Capture (REDCap), a free and secure Health Insurance Portability and Accountability Act–compliant system for managing web-based surveys and databases. Following a brief description of the study, the screener included questions about potential participants’ age, race, ethnicity, sexual identity, sex assigned at birth, gender (including transgender status), high school name, city and state of their high school, computer access, camera access, audio access, microphone access, and contact
information. All participants received the ARA and SGMY resource lists after the screener.

**Purposive Sampling**

After potential participants completed the screener survey, a research assistant will assess their eligibility. The research assistant will send a sociodemographically diverse group of eligible youth a web-based consent form, which must be completed to participate.

**Consent Process**

Potential participants are sent a link to a web-based consent form administered via DocuSign, a website that allows participants to safely and securely use a virtual signature. Youth in this study will consent for themselves because we obtained a waiver of parental consent. Our study is no more than minimal risk and requiring SGMY to obtain parental permission could out them as SGMY to their parents or guardians, which may put them at an increased risk of experiencing abuse. The consent form described all essential components of the study, including (but not limited to) the study purpose, the study background, study risks and benefits, privacy and confidentiality, participant payments, and voluntary nature of the study.

Once the consent form is virtually signed, the research assistant receives a PDF version of the signed form and emails or texts participants a link to the Zoom videoconference meeting, where the web-based human-centered design sessions are conducted. The message includes instructions on how to best prepare for the session and how to access MURAL, the web-based collaborative workspace used during our web-based sessions. The participants are sent reminders 2-3 times before each session.

In addition to the web-based consent form, participants provide verbal consent at the beginning of each web-based human-centered design session. A research assistant reads aloud a verbal consent script, asks if there are any questions, and then participants provide their consent by using Zoom’s thumbs up or via Zoom’s chat feature. In addition, before completing the follow-up survey, a brief consent script is provided to participants and participants voluntarily consented to taking the survey using a click-to-consent procedure.

**Human-Centered Design Activities**

We use MURAL and Zoom to conduct the web-based human-centered design activities. With each cohort of SGMY, we will conduct 4 sessions, each lasting 1.5 hours in length. All the sessions will be conducted in English. We will audio-record each session and take pictures of each session’s activity results. We only record the participants’ voices and the resultant data (not images of participants). The participants will receive a US $25 incentive for each session. Participants do not have to attend all sessions and can begin at any time, although we encourage attendance at all sessions because output from some sessions are then used as inputs in subsequent sessions.

Each session begins with an introduction to Zoom, an icebreaker, and (in the first 2 sessions) a short lesson on the topic at hand (ie, ARA in SGMY). Next, the participants are randomly assigned to different Zoom breakout rooms, with each room composed of 2 to 5 participants with 1 or 2 facilitators each. In these small groups, the facilitators introduce participants to MURAL and guide the participants through a series of human-centered design activities. Each session’s human-centered design activities are outlined and briefly described in Table 1 and described in detail in the sections below. At the end of each meeting, all participants are brought back together and report on the ideas generated during the human-centered design activities.
### Table 1. Human-centered design activities by session.

<table>
<thead>
<tr>
<th>Session and activity</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rose, thorn, bud</td>
<td>Rose, thorn, bud</td>
<td>Creative matrix</td>
<td>Round robin</td>
</tr>
<tr>
<td>Purpose</td>
<td>To have participants individually brainstorm healthy, unhealthy, and questionable aspects of intimate relationships for SGMY&lt;sup&gt;a&lt;/sup&gt;</td>
<td>To have participants provide feedback on the original SHARP&lt;sup&gt;b&lt;/sup&gt; materials or the adapted SGMY-specific SHARP materials</td>
<td>To have participants brainstorm intervention ideas at each level of the social-ecological model for reducing SGMY inequities in adolescent relationship abuse</td>
<td>To have participants evolve their intervention ideas into fuller intervention concepts using quick drafting and iteration via group authorship</td>
</tr>
<tr>
<td></td>
<td>Affinity clustering</td>
<td>Affinity clustering</td>
<td>Impact-difficulty matrix</td>
<td>Visualize the vote</td>
</tr>
<tr>
<td>Purpose</td>
<td>To have participants discuss all the healthy, unhealthy, and questionable aspects of intimate relationships they generated, and to have participants group similar ideas together</td>
<td>To have participants discuss all their feedback on the SHARP-related materials and to have participants group similar ideas together</td>
<td>To have participants plot their self-generated intervention ideas based on their potential ease of implementation and potential impact, thereby prioritizing the intervention ideas with the lowest potential resource expenditure and greatest potential impact</td>
<td>To quickly poll SGMY’s preferences and opinions about two of their favorite intervention concepts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Concept poster</td>
</tr>
<tr>
<td>Purpose</td>
<td></td>
<td></td>
<td></td>
<td>To have participants work together to refine intervention concepts by illustrating and describing its essential elements</td>
</tr>
</tbody>
</table>

<sup>a</sup>SGMY: sexual and gender minority youth.

<sup>b</sup>SHARP: School Health Center Healthy Adolescent Relationships Program.

**Session 1**

The purpose of our first session is to elucidate what SGMY believe are healthy and unhealthy characteristics of intimate relationships (study aim 2). We will accomplish this by using 2 human-centered design techniques: rose-thorn-bud and affinity clustering activities. Both activities are set up in a single MURAL workspace (Figure 2).
Rose-thorn-bud is a technique in which participants identify different aspects of a concept [20,21], in this case, healthy intimate relationships. Each participant is assigned a virtual desk (Figure 2) with 3 different colored sticky notes: pink, blue, and green. The facilitator explains the color-coding system to the participants. Participants are to type aspects of healthy relationships on pink sticky notes (or roses); they are then instructed to type components or aspects of unhealthy relationships on blue sticky notes (or thorns); and green sticky notes (or buds) are used to note aspects of relationships that are uncertain qualities with potential to be healthy and/or unhealthy.

Affinity clustering is used after completing the rose-thorn-bud activity to sort sticky notes according to their similarities and differences [20,21]. The facilitator asks a participant to read and explain one of their sticky notes (of any color) and the facilitator places it along the horizontal axis. The facilitator then asks another participant to describe one of their sticky notes and identify whether it is a similar or unique idea compared with the previously placed sticky note. Once determined, the facilitator then places this sticky note on the workspace below the previous sticky note if similar, or next to the previous sticky note if unique. This process is repeated until all sticky notes are in the workspace and are grouped according to the participants’ satisfaction. It is important to note that during affinity clustering, the facilitator encourages participants to align ideas based on its content, as opposed to the color of the sticky note. This type of clustering allows participants (as well as analysts who later view or interpret the data) to obtain 2 layers of meaning: the overall concepts of the groupings as well as what color sticky notes comprise those groupings. Overall, session 1 yields a visual representation of the SGMY’s mental models of intimate relationships.
Session 2

The purpose of session 2 is to acquire feedback from SGMY about the SHARP intervention materials as well as the adapted SGMY-specific SHARP materials (study aim 3) [48-51]. SHARP is a universal (not SGMY-specific) provider-based intervention implemented in routine school-based health center visits [48]. During clinic visits, SHARP providers introduce a palm-size brochure that contains information about healthy relationships and ARA resources, conduct ARA assessments, make referrals to ARA services, if necessary, and discuss healthy and unhealthy relationships with their patients. All SHARP materials were developed with input from numerous stakeholders, including clinicians, advocates, researchers, and youth. Compared with usual care, SHARP improves adolescents’ recognition of ARA, knowledge of ARA resources, and self-efficacy in using ARA harm reduction behaviors [48]. After the universal SHARP materials were created, investigators and stakeholders (including youth) generated SGMY-specific brochure materials tailored to SGMY [50,51]. We selected the SHARP intervention because (1) it is an illustrative example of an efficacious ARA prevention intervention and study participants may not be familiar with ARA interventions and (2) SGMY can provide feedback about the evidence-based intervention, which might confirm the applicability of the materials to SGMY in 2020 or offer ways to improve the brochure for current clinical practice or future research. We will have SGMY critique the SHARP brochure materials using two human-centered design techniques: rose-thorn-bud and affinity clustering activities. Both of these are set up in a single MURAL workspace (Figures 3 and 4).

Figure 3. MURAL Workspace Setup for Session 2 on Original School Health Center Healthy Adolescent Relationships Program (SHARP) Intervention.
Rose-thorn-bud is a technique for having participants characterize different aspects of a concept (in this case, the SHARP intervention) as good, problematic, or having potential [20,21]. Each participant is assigned a virtual desk (Figures 3 and 4) with 3 different colored sticky notes: pink, blue, and green. The facilitator explained the color-coding system to the participants. Participants are asked to write the components and aspects of what they like about the SGMY intervention on pink sticky notes (or roses); they are then instructed to type components they dislike on blue sticky notes (or thorns); and green sticky notes (or buds) are used to note aspects that have the potential to be inclusive but could be improved. The facilitator shows participants how to create additional sticky notes and instructs the participants to write one idea per sticky note, generating as many sticky notes as possible. For 7 minutes, participants work independently to complete their sticky notes.

Affinity clustering is used after completing the rose-thorn-bud activity to sort items according to their similarities and differences [20,21]. The facilitator asks one participant to read and explain one of their sticky notes (of any color) and the facilitator places it along the horizontal axis. The facilitator then asks another participant to describe one of their sticky notes and identify whether it is a similar or unique idea compared with the previously placed sticky note. Once determined, the facilitator then places this sticky note on the workspace below the previous sticky note if similar, or next to the previous sticky note if unique. This process is repeated until all sticky notes are in the workspace and are grouped according to the participants’ satisfaction.

**Session 3**

The purpose of session 3 is to brainstorm novel intervention components to reduce ARA inequities for SGMY and force-rank the intervention components based on their potential ease of implementation and potential impact (study aim 4). We accomplish this by using 2 human-centered design strategies: a creative matrix and an importance-difficulty matrix. All of these are conducted in one MURAL workspace (Figure 5).
The creative matrix is used to brainstorm as many intervention concepts for reducing ARA among SGMY as possible by leveraging the power of constraints (column by column) and a combination of goals (column headings) and enablers (essentially categories of potential solutions) [16,18,20,21,23,26,27]. As shown in Figure 5, the column headers indicate different levels of the social-ecological model, that is, individual, interpersonal, family and community, structural, and organizational determinants, whereas the row headings indicate the modalities: technology, events or programs, physical environments, policies or procedures, games or competitions, and wild card (i.e., any other ideas that do not fit with the labeled categories). By providing different levels of the social-ecological model, SGMY are primed to think about creating intervention components that address known risk factors (e.g., minority stressors [57]) and protective factors that contribute to ARA inequities for SGMY [57-63]. The activity is completed in 1 column at a time, with 4 minutes per column. Participants...
are asked to develop ideas (one idea per sticky note) for each cell or box by working independently. Participants are working individually at this time, reducing groupthink (eg, influence of more vocal contributors). In this activity, participants are encouraged to generate as many ideas as possible within a short period. Emphasis is typically placed on quantity (over quality) to eliminate barriers to contribution, get participants comfortable with the activity, and give them the opportunity to share ideas they may have already had before being prompted with the matrix and create space for new ideas informed by exploration of each row and column combination.

After completing the creative matrix, participants engage in a structured activity to funnel the large number of ideas generated in the creative matrix into a manageable number of ideas to bring to the next activity. Facilitators ask participants to identify 4 sticky notes each: most important, most doable, their favorite, and the most out-of-the-box. These 4 criteria for down-selection were designed to ensure that ideas pursued in subsequent activities are sufficiently diverse. Each participant works independently for 3 minutes while they choose their ideas. The participants cannot choose the same ideas.

An importance difficul-tty matrix [20-22,26] is then used to force-rank the ideas identified in the previous step. First, participants are asked to collaboratively rank the ideas that they chose on the level of importance along the x-axis. Importance here is defined as how important or impactful the participants think the idea is to reduce ARA for SGMY. Next, they are asked to rank these ideas on the level of difficulty along the y-axis while retaining the same ordered ranking of likely impact or importance. The purpose of this activity is to identify the intervention concepts that are the most important to the participants and, given the inevitable resource constraints, account for their potential difficulty to implement.

Session 4
The purpose of session 4 is to generate, iterate, vote on, and refine intervention concepts to reduce ARA inequities for SGMY (study aim 5). We accomplish this by using three human-centered design strategies: round robin, visualize the vote, and concept posters. All of these are set up in one MURAL workspace (Figure 6).
In total, 5 of the potentially most impactful and easiest to design and implement ideas generated from the importance-difficulty matrix in session 3 are carried over to session 4 for the round robin activity. Round robin is used to evolve ideas into fuller intervention concepts using drafting and iteration among participants [20,21,24,25]. Each desk has 4 sections: problem statement (idea from previous session), proposed solution, why the solution might fail, and the final concept. This structure was developed by the LUMA Institute (see Figure 7 for a reader-friendly version of the round robin structure) [20,21]. Each participant starts at their own desk (a specific rectangle in the MURAL workspace) with the initial intervention concept. For 3 minutes, participants are asked to propose a solution to their intervention concept. Then, participants rotate workspaces by moving to the space to their right (or, if at the last desk, moving to the first desk). Then, participants are asked to provide feedback on why the next intervention concept might fail for 5 minutes. Participants move to the desk to their right one final
Visualizing the vote is an activity that allows the facilitator to anonymously poll the participants and is a feature offered by MURAL [20,21]. After reading through the final round robin concepts, facilitators show participants where the voting ballots are and ask participants to vote for their 2 favorite concepts. Once this voting process is complete, 2 round robin concepts with the most votes overall are carried over to the final activity, the concept poster.

The concept poster is used to illustrate and present the main points of the new intervention concepts [20,21]. The concept poster format, developed by the LUMA Institute [20,21], includes generating a creative title and identifying the following: the target population, the problem it solves, the big idea, how it works, why it might fail, what a prototype looks like, and how to measure success (see Figure 8 for a reader-friendly version of the concept poster activity). Facilitators first show the participants an example of a completed concept poster to illustrate the purpose of the concept poster. Participants are then invited to collaboratively develop 2 concept posters (allowing 15 minutes per poster). During this activity, facilitators answer participants’ questions and provide guidance on an as-needed basis.
Figure 8. Example of one concept poster setup in MURAL.

Concept poster

What is the concept called?

Type your concept name here...

Who is it for? What problem does it solve? What is the big idea?

Illustrate how it works!

Why might it fail? What should we prototype and test? How might we measure success?

Follow-Up Survey Data Collection

After session 4, participants are asked to complete self-administered follow-up surveys via REDCap. The follow-up surveys assess participants’ perceptions about the feasibility, appropriateness, and acceptability of the web-based human-centered design session, as well as participants’ qualitative feedback on session logistics. The follow-up survey is activated and sent within 1 week after session 4. Surveys remain open for up to 2 weeks. The follow-up survey contains 3 pages, with a mean of 10 items per page (SD 6; range 6-17). While completing each survey, participants were able to change their answers by clicking a Back button. An incentive of US$10 was given after completion of the follow-up survey.

Follow-Up Survey Measures

To accomplish aim 1, we assess participants’ perceptions of the feasibility, acceptability, and appropriateness of the web-based human-centered design sessions.

Feasibility is defined as the perception among participants that human-centered design sessions can be successfully implemented or carried out on the web [64]. Feasibility is measured using the Feasibility of Intervention Measure [64-66], which is a valid and internally consistent scale based on 4 positively worded items (eg, “The online sessions for this project were implementable”) that use a 5-point Likert scale ranging from completely disagree (1) to completely agree (5). We will create a mean across all items, with a higher mean score indicating greater feasibility.

Acceptability is defined as the perception among stakeholders that web-based human-centered design sessions are satisfactory or agreeable [66]. Acceptability is measured using the Acceptability of Intervention Measure [65-67], which is a valid and internally consistent scale based on 4 positively worded items (eg, “The online sessions for this project seemed doable”) that use a 5-point Likert scale ranging from completely disagree (1) to completely agree (5). We will create a mean across all items, with a higher mean score indicating greater acceptability.

Appropriateness is defined as the perception among participants that web-based human-centered design sessions are relevant, compatible, and suitable to address a specific issue [66].
Appropriateness is measured using the Intervention Appropriateness Measure [65], which is a valid and internally consistent scale based on 4 positively worded items (eg, “The online sessions for this project seemed applicable”) that use a 5-point Likert scale ranging from completely disagree (1) to completely agree (5). We will create a mean across all items, with a higher mean score indicating greater appropriateness.

In addition to our standardized measures, we ask participants exploratory open-ended questions including about what they liked most and liked the least about the web-based suggestions as well as ways to improve the sessions and retention. We also asked closed-ended questions about their opinions about session length and frequency.

Data Analyses
For quantitative data analysis, we will use StataSE version 15 (StataCorp). For qualitative data analysis, we will transcribe, de-identify, and quality check all audio-recorded data from sessions 1 to 4 [68-71]. We will also export all session data from MURAL into text-based documents. We will then perform qualitative coding in Dedoose software [72], cross-referencing the audio transcriptions with PDF exports of the MURAL workspaces.

Sample Characteristics
Using data from the web-based screener survey, we will use frequencies and percentages to describe the demographics of our study sample.

Analyses for Study Aim 1
To assess the feasibility, acceptability, and appropriateness of conducting web-based human-centered design sessions with SGMY, we will calculate the means and 95% CIs for our primary outcome measures of feasibility, acceptability, and appropriateness. Our a priori benchmark for the success of meeting this hypothesis is obtaining a mean significantly higher than 3.75 (out of 5.00).

Analyses for Study Aims 2 to 5
To elucidate what SGMY believe are healthy and unhealthy characteristics of intimate relationships, we will use content analyses [73]. Using data from each session separately, we will code the artifact or audio-recording transcripts using an open-coding approach (ie, we will not create an a priori codebook, but will use in vivo coding instead) [68,69,74]. A total of 2 qualitative coders will independently read data from that session. Coders will convene to discuss preliminary findings and develop a draft codebook using inductive coding, allowing new codes to be included in the codebook as they emerge. Once the coders agree that all the major codes are identified, we will create a final codebook with definitions, rules, and examples for each code [70,71]. The 2 coders will then recode all data using the final codes. We will calculate the interrater reliability (ie, Kappa statistic) to examine code application between coders [75]. Coders will discuss any discrepancies until they reach an agreement; any disagreements will be discussed and resolved during research team meetings. We will use either a qualitative descriptive coding approach [76] (wherein we describe and count the number of code applications) or axial coding [77] (wherein we combine inductive codes into broader categories to define emerging patterns or themes). We will identify and interpret patterns in the data using thematic analysis, as informed by Braun and Clark [78].

Sample Size and Power Calculation
We calculate statistical power based on our primary outcomes (study aim 1), per the best practices for feasibility studies [64,79-84]. Given a conservative sample size of 45 participants and 5% type I error rate, we have the ability to estimate a 95% CI margin of error ≤0.43 for our primary outcomes, which we derived from the largest upper 95% CI limit of the SD measures in previous studies that used the same outcome measures [16-27].

For our other study aims (study aims 2-5) about the outcomes of our human-centered design methods, the purpose is idea generation, not saturation [16-27]. Given our previous work implementing similar activities, a sample of 45 to 60 participants is sufficient to generate unique and useful ideas.

Researcher Characteristics and Reflexivity
Our team represents a wide range of sexual identities (ie, gay, lesbian, queer, bisexual, and heterosexual) and gender identities (ie, nonbinary, cisgender women, and cisgender men). We also comprise a multidisciplinary multitiered team, representing the fields of public health, counseling, human-centered design, library science, medicine, medical anthropology, political science, psychology, rehabilitation science, social work, and many roles in academia, including undergraduate, masters, research assistants, assistant professors, and full professors. Our diverse representation of sexual and gender identities complements our range of multidisciplinary expertise and professional backgrounds in informing data collection and analysis through lived experience.

We acknowledge that there are several ways in which our backgrounds may have influenced our research. Primarily, our facilitators for the human-centered design sessions were sexual minorities and gender minorities. This could help put our SGMY participants at ease, especially in rare instances when the interviewer disclosed their sexual or gender minority identities to participants. Furthermore, the diversity and range of disciplines represented by our research team provide us with the ability to interpret our findings.

Ethics Statement
All study procedures were approved by the Human Research Protection Office of the University of Pittsburgh. To protect participants from having to reveal their sexual or gender minority identities to their caregivers, thereby potentially putting them in harm’s way, we received a waiver of parental consent [85]. This allowed the participants to provide consent for themselves. To further protect participants, we asked SGMY (before and during each session) to find a quiet and private space where they could participate in activities with minimal distraction and interruption. We also allowed participants to use Zoom’s chat function if they are unable or feel uncomfortable speaking out loud. This study was also protected by a Certificate of Confidentiality from the National Institutes
of Health. It is important to note that we never inquired about participants’ personal experiences with ARA. Nevertheless, we provided all people who completed the screener with the ARA and SGMY resource lists. In addition, the study participants received resource lists after each session.

**Results**

This study was funded in February 2020. Data collection began in August 2020 and will be completed in April 2021. From August 2020 through December 2020, 778 individuals clicked the link to the screening questionnaire (Figure 9). In total, 370 individuals completed the screening questionnaire, of which 274 individuals met all eligibility criteria. A total of 50 participants were invited to participate, and 22 consented to participate. In total, 16 SGMY participated in at least one human-centered design session. All data collection will be completed by April 2021.

**Discussion**

**Impact**

This study protocol has several methodological innovations that can inform future public health research that aims to incorporate methods from the field of human-centered design. First, our method is an example of how human-centered design research can be performed on the web. This is important because it is a resource-friendly and accessible method for engaging diverse stakeholders from a wide geographic region. Moreover, the
COVID-19 pandemic hindered our ability to engage safely with participants, and these web-based methods enable our research to continue despite current limitations on in-person research activities. Our study also demonstrates how to combine human-centered design, which is often quite flexible, into research processes, which tend to be much more structured. By rigorously testing the feasibility of such an approach, our study has the potential to demonstrate and codify the use of human-centered design as a novel stakeholder-engaged research method.

Our study has the potential to lead to many substantive innovations in the field of ARA interventions among SGMY. If successful, our study could yield several novel intervention concepts. Importantly, these interventions are directly derived from SGMY themselves, as opposed to researchers. By centering youth voices and opinions in this manner, the generated interventions may be highly acceptable and impactful, but that cannot be determined until the interventions are tested further. Nevertheless, given the great lack of interventions, our study can help catalyze the field of SGMY ARA intervention research.

**Limitations**

Although there are many strengths to this study, it is not without limitations. Although participants in our sample are sociodemographically diverse, they are not necessarily representative. For example, SGMY without internet access are excluded from our study, and sessions were only conducted in English. In addition to youth, the inclusion of other stakeholders (eg, parents and school personnel) would likely have important insights for intervention concepts, but they are not included in this study. Despite our study yielding potentially new intervention concepts, this study will not produce complete interventions. Additional work from researchers, designers, and stakeholders will be necessary to develop and test the derived interventions. Finally, this is a pilot study testing the feasibility of our methods. We are not testing the effectiveness of our methods versus more traditional methods (eg, focus groups) for producing intervention concepts, which can be executed in future trials if our methodology proves to be feasible.

**Conclusions**

Interventions to reduce ARA among SGMY are lacking. To address this gap, our study investigates the feasibility of a new method for generating new intervention concepts. This work has the potential to contribute to substantive health impacts as well as immediate methodological impact by integrating human-centered design methods into public health research.

**Acknowledgments**

This research was supported by the National Institutes of Health through the National Center for Advancing Translational Sciences (grant UL1TR001857) and the National Institute on Alcohol Abuse and Alcoholism (grant K01AA027564 to RC). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The National Institutes of Health was not involved in the study design, writing of the protocol, or the decision to submit for publication.

**Conflicts of Interest**

None declared.

**References**


**Abbreviations**

**ARA:** adolescent relationship abuse  
**REDCap:** Research Electronic Data Capture  
**SGMY:** sexual and gender minority youth  
**SHARP:** School Health Center Healthy Adolescent Relationships Program

©Robert W S Coulter, Shannon Mitchell, Kelly Prangley, Seth Smallwood, Leyna Bonanno, Elizabeth N Foster, Abby Wilson, Elizabeth Miller, Carla D Chugani. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 12.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Pregnancy Outcomes and Child Development Effects of SARS-CoV-2 Infection (PROUDEST Trial): Protocol for a Multicenter, Prospective Cohort Study

Geraldo Magela Fernandes1*, MSc, MD; Felipe Motta1*, MD; Lizzandra Moura Paravidine Sasak1*, MSc, MD; Ângelo Pereira Da Silva2, MD; Andreza Monforte Miranda3, MSc; Aleida Oliveira De Carvalho2, BSc; Ana Paula Monteiro Gomides4, MD, PhD; Alexandre Anderson De Sousa Munhoz Soares1, MD, PhD; Agenor De Castro Moreira Dos Santos Jr5, BSc; Caroline De Oliveira Alves1, MSc, PhD; Ciro Martins Gomes1, MD, PhD; Clara Correia De Siracusa1, MD; David Alves De Araujo Jr2, MD; Dayde Lane Mendonça-Silva2, MSc, PhD; José Alfredo Lacerda De Jesus1, MSc, MD, PhD; Karina Nascimento Costa1, MSc, MD, PhD; Maria Eduarda Canellas De Castro2, MD; Patrícia Shu Kurizky2, MD, PhD; Paulo Sérgio França1, MSc, MD, PhD; Rosana Tristão1, MSc, PhD; Yacara Ribeiro Pereira2, BSc; Luiz Claudio Gonçalves De Castro1, MSc, MD, PhD; Alberto Moreno Zaconeta1, MSc, MD, PhD; Cleanderson Pires De Albuquerque2, MSc; Licia Maria Henrique Da Mota1, MD, PhD

1Faculty of Medicine, University of Brasília, Brasília - DF, Brazil
2Department of Rheumatology, University Hospital of Brasília, Brasília, Brazil
3State Health Department of the Federal District, Brasília, Brazil
4Faculty of Medicine, University Center of Brasília, Brasília, Brazil
5Central Laboratory of Public Health of the Federal District, Brasília, Brazil
*these authors contributed equally

Corresponding Author:
Geraldo Magela Fernandes, MSc, MD
Faculty of Medicine
University of Brasília
Campus Universitário Darcy Ribeiro, UNB Área 1 - Asa Norte
Brasília - DF, 70910-900
Brazil
Phone: 55 61981070071
Email: geraldomafer@gmail.com

Abstract

Background: A growing body of evidence suggests that SARS-COV-2 infection during pregnancy may affect maternal-fetal outcomes and possibly result in implications for the long-term development of SARS-CoV-2–exposed children.

Objective: The PROUDEST (Pregnancy Outcomes and Child Development Effects of SARS-CoV-2 Infection Study) is a multicenter, prospective cohort study designed to elucidate the repercussions of COVID-19 for the global health of mothers and their children.

Methods: The PROUDEST trial comprises 2 prospective, sequential substudies. The PREGNANT substudy will clinically assess the effects of SARS-CoV-2 infection on pregnancy, childbirth, and puerperium from a mechanistic standpoint to elucidate the pregnancy-related inflammatory and immunological phenomena underlying COVID-19. Pregnant women aged 18-40 years who have been exposed (proven with laboratory tests) to SARS-CoV-2 (group A; n=300) will be compared to control subjects with no laboratory evidence of in-pregnancy exposure to the virus (group B; n=300). Subjects exposed to other infections during pregnancy will be excluded. The BORN substudy is a long-term follow-up study that will assess the offspring of women who enrolled in the prior substudy. It will describe the effects of SARS-CoV-2 exposure during pregnancy on children’s growth, neurodevelopment, and metabolism from birth up to 5 years of age. It includes two comparison groups; group A (exposed; n=300) comprises children born from SARS-CoV-2–exposed pregnancies, and group B (controls; n=300) comprises children born from nonexposed mothers.
Results: Recruitment began in July 2020, and as of January 2021, 260 pregnant women who were infected with SARS-CoV-2 during pregnancy and 160 newborns have been included in the study. Data analysis is scheduled to start after all data are collected.

Conclusions: Upon completion of the study, we expect to have comprehensive data that will provide a better understanding of the effects of SARS-CoV-2 infection and related inflammatory and immunological processes on pregnancy, puerperium, and infancy. Our findings will inform clinical decisions regarding the care of SARS-CoV-2–exposed mothers and children and support the development of evidence-based public health policies.

Trial Registration: Brazilian Register of Clinical Trials RBR65QXS2; https://ensaiosclinicos.gov.br/rg/RBR-65qxs2

International Registered Report Identifier (IRRID): DERR1-10.2196/26477

(JMIR Res Protoc 2021;10(4):e26477) doi:10.2196/26477

KEYWORDS
SARS-CoV-2; COVID-19; pregnancy; neonate; children; outcome; development; prospective; cohort; women; fetus; baby; implication

Introduction

Background

The natural history of COVID-19, a disease caused by SARS-CoV-2, is still being written. Individuals infected with SARS-CoV-2 may present with a broad spectrum of clinical manifestations, from no symptoms to dramatically progressive disease symptoms that eventually result in death [1,2]. Some individuals develop intense inflammatory and procoagulant responses that can result in severe pulmonary damage, which is the main cause of COVID-19 morbidity and mortality.

The pathophysiological phenomena that take place in other human tissues that are potentially targeted by SARS-CoV-2 need further investigation. A central nervous system viral tropism has been postulated based on reports of neurological events such as stroke, acute hemorrhagic encephalopathy, seizures, and the loss of smell and taste [3,4].

To date, little is known about the effects of COVID-19 on women during pregnancy and puerperium [5,6] and its consequences for women’s offspring (from the neonatal period through the first years of life) [7-9]. Thus, there is still a lack of a robust evidence base for the proper management of these mothers and children.

Data from previous epidemics of viral-induced respiratory distress syndrome have shown that pregnant and puerperal women have a high risk of developing life-threatening clinical outcomes [10]. These generally worse outcomes have been attributed to physiological changes in the immune and cardiopulmonary systems that occur during pregnancy [11]. Examples of epidemics include the H1N1 influenza, SARS-CoV (severe acute respiratory syndrome coronavirus), and MERS-CoV (Middle East respiratory syndrome coronavirus) epidemics, during which pregnancy mortality rates reached 27% [10,12,13].

Despite the structural similarities between coronaviruses, the initial reports for pregnant women with COVID-19 showed lower rates of intensive care unit admission, orotracheal intubation, and death during the SARS-CoV-2 outbreak than those during the SARS-CoV and MERS-CoV outbreaks [9,13]. However, more recent papers have shown higher morbidity and mortality rates among pregnant women than those among nonpregnant women. Several aspects of embryo implantation [14], placental development [15], and delivery dynamics [16] seem to be impaired by the inflammatory response driven by immune cell subtypes at the maternal-fetal interface [17]. Such phenomena may precipitate preeclampsia, spontaneous abortion, intrauterine growth restriction, and premature birth [18-21].

Most of the available data on pregnant women exposed to SARS-CoV-2 were obtained during the second half of pregnancy. Thus, SARS-CoV-2 infection during all stages of pregnancy, including the early stages of gestation, has not been fully investigated. Nevertheless, as the disease spreads worldwide, more women are being exposed to the virus during early gestation and midgestation, and new data have been accruing [22].

Studies evaluating the vertical transmission of SARS-CoV-2 are still inconclusive [6,23-25]. Investigations of placentas from women infected with SARS-CoV-2 have suggested that there is a low likelihood of viral transplacental transmission. However, the potentially hazardous effects of inflammatory and prothrombotic environments on placental function and, consequently, fetal growth could not be ruled out [26-28].

To date, the few reports on postnatally infected neonates have shown that they exhibit either no symptoms or mild clinical forms of COVID-19 with favorable outcomes. However, the younger the infant is, the higher the risk of critical outcomes [22,29].

Maternal SARS-CoV-2 infection would potentially expose a fetus not only to direct viral effects but also to the placental inflammatory response and the maternal cytokine storm [5,7,22]. Such processes and their consequences have not been extensively studied. The understanding of these phenomena should contribute to the proper management of children born to mothers infected with SARS-CoV-2.

A case series on the clinical aspects of newborns of COVID-19–exposed mothers reported a low risk of adverse outcomes for late pregnancy exposure and stated that there is a paramount need for close follow-ups [30]. Other reports have presented data showing no adverse effects on neonates born to mothers who tested positive for COVID-19. Furthermore, Liu et al [31] described 19 completely asymptomatic neonates from Wuhan, China.
A few studies however have reported that SARS-CoV-2 test–negative neonates born to mothers who tested positive and developed critical illnesses might present ongoing clinical profiles. This is suggestive of the potential impact of inflammatory processes on fetal physiology. Romagano et al [32] reported a prevalence rate of 6.9% for symptomatic pregnant women infected with SARS-CoV-2 among 1053 deliveries at a large hospital network in New Jersey, United States. They reported that 8 pregnant women were critically ill and 7 neonates tested negative (via reverse transcriptase–polymerase chain reaction [RT-PCR]); 1 neonate was not yet delivered at the time of testing. All neonates were preterm and appropriately sized for their gestational age except for one (small for their gestational age). They were all separated from their mothers after delivery, and all of them developed respiratory distress and required neonatal intensive care unit admission. Anemia and hyperbilirubinemia of prematurity, temperature instability, and feeding problems were reported in some of the neonates.

Several other studies have reported symptoms among SARS-CoV-2 test–negative neonates born to mothers with COVID-19, such as rashes [33], facial ulceration [33], the need for noninvasive oxygen support [33], transient lymphocytopenia [34], impaired liver function [34], disseminated intravascular coagulation, and even multiple organ failure leading to death [35]. There are many critical questions regarding the standards of care for SARS-CoV-2–exposed pregnant women and their offspring that have yet to be answered, and guidelines are still being developed around the world [8,36,37]. Therefore, the overall purpose of this study is to describe the effects of in-pregnancy SARS-CoV-2 infection and related inflammatory and immunological phenomena on the health of SARS-CoV-2–exposed women and their offspring.

Objective

Our specific aims are (1) to study the effects of COVID-19 on maternal and obstetric morbidity and mortality, including those of indicators such as preeclampsia, abortion, fetal malformation, fetal growth, and premature birth; (2) to investigate the presence of SARS-CoV-2 and anti–SARS-CoV-2 antibodies in the cerebrospinal fluid (CSF) of women with symptomatic COVID-19 undergoing spinal anesthesia for a cesarean section; (3) to determine the serum proinflammatory and regulatory cytokine profiles of pregnant women with symptomatic COVID-19; (4) to determine the CSF proinflammatory and regulatory cytokine profiles of pregnant women with symptomatic COVID-19 undergoing spinal anesthesia for a cesarean section; (5) to study the histopathological markers of inflammatory and thrombotic phenomena in the placenta; (6) to study the correlations between the aforementioned serum and histologic biomarkers of COVID-19 and the outcomes of pregnancy, delivery, puerperium, and childbirth as well as the correlations between biomarkers and short- and long-term health outcomes during infancy; (7) to study the association between the use of maternal pharmacological therapy for treating COVID-19 and offsprings’ health outcomes; (8) to evaluate the effects of COVID-19 exposure during different stages of pregnancy on fetal, neonatal, and infantile morbidity and mortality; and (9) to evaluate the effects of in-pregnancy COVID-19 exposure on children’s somatic and neurological development and energy metabolism from birth up to 5 years of age.

Methods

Study Design

The PROUDEST (Pregnancy Outcomes and Child Development Effects of SARS-CoV-2 Infection Study) is a multicenter, longitudinal, prospective observational study that will be conducted in two sequential stages—the PREGNANT and BORN branches (or substudies). Each stage will have two parallel groups (exposed and nonexposed) for comparisons. The PROUDEST is designed to address the multifaceted questions surrounding the impact of COVID-19 exposure during pregnancy on the global health of mother-child dyads (Figure 1).

Figure 1. Pregnancy Outcomes and Child Development Effects of SARS-CoV-2 Infection Study follow-up flowchart.

The PREGNANT substudy will follow—until day 21 postpartum—pregnant women who are exposed to SARS-CoV-2 at any phase of gestation and compare them to a control group consisting of nonexposed pregnant women. The BORN substudy will follow the children of the women included in the preceding (PREGNANT) branch. These children will be allocated into two comparison groups (exposed and nonexposed) according to their mothers’ in-pregnancy exposure status and will be followed up by a multidisciplinary team of health professionals from birth up to the age of 5 years. This team will conduct regular consultations every month up until the children reach 6 months of age, every 3 months up until the children reach 2 years of age, and every 6 months up until the children reach 5 years of age. Mothers and children may attend nonscheduled visits, as needed, for clinical reasons as well as specific appointments for conducting the procedures and tests described in this protocol.
The PROUDEST will be conducted from July 2020 to December 2026 in Brasilia, Brazil. The recruitment of pregnant and newborn dyads will be carried out by using data from the Epidemiological Surveillance Center of the Federal District. These mother-fetus dyads will be followed up to childbirth (and puerperium in the case of mothers) until December 2021, which is when the last included dyads are expected to undergo childbirth at two different hospitals—the University Hospital of Brasilia and Asa Norte Regional Hospital (the reference public medical center for COVID-19 in the Federal District in Brazil). Both hospitals are located in central urban areas and are included in the Brazilian public health system (Sistema Único de Saúde), which primarily serves the low-income population. Thus, the results of the PREGNANT substudy will be published as soon as the analyses are completed. The children will be followed from childbirth up to December 2026, which is when the last admitted neonate will turn 5 years old. As the BORN substudy is lengthy, partial results may be disclosed during the course of the study, but the final data set will be made available in the second half of 2026.

Pregnant women included in the study must be aged >18 years. COVID-19 exposure will be defined as a first-time RT-PCR test, serology test, or rapid test that returns positive results during pregnancy and is confirmed by a second test. Nonexposure to COVID-19 will be defined as asymptomatic pregnant women with negative serology tests (immunoglobulin G [IgG] and immunoglobulin M [IgM] tests), which will be conducted at 14-21 days postpartum.

Pregnant women with preexisting chronic diseases (except diabetes and hypertension); those taking continuous medications; those who consume tobacco, alcohol, or other drugs; and those with other suspected or confirmed congenital infections will be excluded.

Neonates whose mothers qualified for inclusion in the PREGNANT substudy (had these women been screened) may also be admitted to the BORN substudy, even if their mothers did not participate in the preceding branch.

Children initially assigned to the control (nonexposed) group who later become infected with SARS-CoV-2 (as confirmed via laboratory tests) during follow-up will be excluded from all analyses (from the time of SARS-CoV-2 infection diagnosis onward). However, they will continue to receive assistance under the same standards until the end of the study.

Sample Size Calculation
No precise data are available on the prevalence of SARS-CoV-2 infection among pregnant women in Brazil, but international reports have estimated that up to 15.3% of all pregnancies have been exposed to the virus [38]. Recent data have indicated a birth rate of 44,195 newborns per year in the Federal District [39]. Thus, after considering an “infinite” population (>20,000 pregnant women), assuming a 15% prevalence of SARS-CoV-2–exposed pregnancies, and accounting for a confidence level of 95% and a margin of error of 5%, the minimum sample size for a random sample of SARS-CoV-2–exposed women would be 195. This calculation expectedly yielded a similar number for a random sample of SARS-CoV-2–exposed children. If we set the expected dropout rate for the BORN substudy to 20%, the required number of SARS-CoV-2–exposed mothers (those giving birth to the BORN participants) would increase to 234.

Our sampling approach however is not truly random; it is based on convenience, as eligible subjects will present to the recruitment centers. The aforementioned calculations only serve as a reference for avoiding the overestimation of the inclusion of participants. Given the limited amount of available knowledge regarding the effects of SARS-CoV-2 infection on pregnancy and child development, which resulted in the eminently exploratory character of our study, we adopted an “as much as feasible, but no more than reasonable” approach for defining the sample size.

We set an a priori number of 300 SARS-CoV-2–exposed women for the PREGNANT phase. This will result in the inclusion of the expected 300 SARS-CoV-2–exposed children in the BORN phase. The subject allocation rate between the exposed and control groups was 1:1. This indicated the need for an additional 300 mothers and 300 children to constitute the nonexposed control groups. Hence, the overall sample size of the PROUDEST was set to 1200 participants (600 mother-child dyads consisting of 300 SARS-CoV-2–exposed mother-child dyads and 300 control dyads).

To promote participant retention and completed follow-ups at pregnancy and pediatric outpatient clinics, we will actively search for patients via phone and email.

Procedures
A host of clinical, psychological, neurodevelopmental, biochemical, histological, and imaging assessments will be conducted in accordance with the PROUDEST protocol (Figure 2).
Prenatal data from both the pregnant women and their fetuses will be gathered during the follow-up at the Pregnancy Outpatient Clinic of the University Hospital of Brasilia. These data will consist of the medical and sociodemographic data of the mothers; gestational age; symptoms, interventions, and outcomes related to COVID-19 (for SARS-CoV-2–exposed participants); congenital infection screening results; hypertensive disorders and other pregnancy-specific morbidities; general health assessments; general physical examinations; routine clinical biochemistry tests; and ultrasound scans. These scans will be performed between gestational weeks 11-13 and gestational day 6, from gestational week 22 to gestational week 24, and on a monthly basis in the third trimester of pregnancy to assess fetal growth and morphology, placental morphology, amniotic fluid volume, and Doppler ultrasound results. Maternal blood will be collected at the first prenatal consultation, regardless of gestational age. Antenatal consultations will occur monthly up to week 34, every 2 weeks between weeks 34 and 36, and then weekly up to delivery. During pregnancy, psychological risk assessments and mental health screens will be performed with the Beck Depression Inventory during the first prenatal consultation. Individual psychological care will be provided to pregnant women who have a Beck Depression Inventory score of >12. Mothers will also be physically and psychologically evaluated between days 7 and 21 postpartum during the PREGNANT substudy and after the BORN substudy.

At childbirth, assessments will be conducted to identify the occurrence of dysfunctional labor and the premature rupture of membranes, the type of birth, and delivery outcomes. We will also conduct physical examinations and classifications of the newborns and anthropometry. The early initiation of breastfeeding, the need for neonatal intensive care, and the type of interventions will also be identified. Maternal blood, CSF from women undergoing spinal anesthesia for a cesarean section, and umbilical cord blood samples will be collected.

CSF will be collected immediately before the infusion of the medicine for spinal anesthesia, which will be injected via a sterile syringe at an average dosage of 0.5 ml. CSF will be
collected in a 4-mL cryotube. Blood samples from mothers and the umbilical cord will be collected in a heparinized tube and centrifuged immediately, and the plasma will be stored in 4-mL sealed cryotubes. CSF and blood samples will be subsequently stored at -80 °C for later analysis. The assessment of blood cell counts, inflammation markers (C-reactive protein and procalcitonin), biochemistry (alanine aminotransferase, aspartate aminotransferase, ferritin, alkaline phosphatase, and lactic dehydrogenase) and SARS-CoV-2 tests (RT-PCR and IgM and IgG antibody tests) will be carried out.

Circulating cytokine levels will be evaluated with the Luminex Bio-Plex Pro Human Cytokine 27 platform (Bio-Rad Laboratories). The cytokine profile assessment will analyze chemokines (CXC motif chemokine ligand [CXCL] 8, CC motif chemokine ligand [CCL] 11, CCL3, CCL4, CCL2, CCL5, and CXC10), proinflammatory cytokines (interleukin [IL]-1β, IL-6, tumor necrosis factor, IL-12p70, interferon γ, IL-17A, and IL-15), regulatory cytokines (IL-1Ra, IL-4, IL-5, IL-9, IL-10, and IL-13), and cell growth factors (IL-2, IL-7, basic fibroblast growth factor, platelet-derived growth factor, vascular endothelial growth factor, granulocyte colony-stimulating factor, and granulocyte-macrophage colony-stimulating factor). All procedures will be performed according to the manufacturer’s recommendations.

Analyses will be performed at the clinical biochemistry laboratories of the hospitals where delivery occurred (blood cell count, C-reactive protein, procalcitonin, and routine biochemistry assessments), the Central Laboratory of the Federal District Department of Health (Laboratório Central de Saúde Pública; SARS-CoV-2 tests), and the University of Brasilia laboratory (cytokine profile assessments).

The placenta will be subjected to fresh histopathological analyses for assessing possible morphological and histological changes that may be associated with SARS-CoV2 infection. Histopathological analysis will be conducted according to the Amsterdam protocol [41].

Peripheral blood and CSF samples from the newborns will only be collected if there is clinical need; this will not be done per the routine research protocol. If such specimens are made available, they will also be subjected to the aforementioned analyses.

All newborns will undergo neonatal screening tests in accordance with the recommendations of the Ministry of Health of Brazil. Five drops of blood will be collected on filter paper for the national neonatal screening program. This blood sample will be collected after 48 hours of life and will be used to screen for the following diseases: phenylketonuria, congenital hypothyroidism, biotinidase deficiency, cystic fibrosis, and congenital adrenal hyperplasia. After 24 hours of life and before discharge, a pulse oximetry test will be performed. Oxygen saturation in the right upper limb and one of the lower limbs will be measured. The pulse oximetry test will be considered normal if oxygen saturation is ≥95% and the difference between the limbs is not ≥3%. Hearing screening will be performed between 36 and 48 hours of life by analyzing otoacoustic emissions.

After hospital discharge, all neonates will be followed up at the Pediatric Outpatient Clinic of the University Hospital of Brasilia. Child growth and neurodevelopment will be assessed at all visits. The first visit will be scheduled to occur on day 15 postpartum. Afterward, visits will be conducted monthly during the first 6 months of life. Thereafter, visits will be scheduled every 3 months until the children reach 12 months of age and every 6 months until the children reach 5 years of age. Nonscheduled visits may occur due to urgent clinical needs. The outpatient clinic staff (a multidisciplinary team) will be composed of pediatricians, psychologists, occupational therapists, speech therapists, physiotherapists, and nurses. The psychological effects of SARS-CoV-2 infection on mothers will be assessed with the Edinburgh Postnatal Depression Scale [42]. This assessment will occur more than once until their children reach 6 months of age. Breastfeeding and weaning patterns, dietary habits, nutritional status, and vaccinal status will be assessed throughout the study.

The assessment of children’s neurodevelopment will be carried out until they reach the 60th month of life. This assessment will analyze cognitive, motor, socioemotional, and language-related aspects and adaptive behavior. The evaluation will be conducted by using the Bayley III Child Development Scale (ie, the version validated for Brazilian infants) [43]. From the age of 2.5 years onward, aspects related to intellectual performance will also be assessed by using the Wechsler Preschool and Primary Intelligence Scale, third edition at 6-month intervals [44].

Central nervous system imaging assessments will be carried out via transcranial ultrasound doppler scans, which will be performed between the 15th and 90th day of the child's life. A brain magnetic resonance imaging scan will be performed if altered cephalic perimeter measures, neurological development delays, or abnormal ultrasound doppler scan findings are identified.

Blood will be collected from SARS-CoV-2–exposed children aged 12 and 24 months to assess their metabolic profiles, which will be used to identify the long-term effects of SARS-CoV-2 infection on systemic metabolism that are potentially driven by past viral exposure and associated inflammatory responses. The examination will consist of assessments for energy metabolism markers (serum lipids, glucose, and insulin), thyroid function markers (thyroid-stimulating hormone and free thyroxine), bone metabolism markers (parathyroid hormone, calcium, phosphate, alkaline phosphatase, and 25-hydroxyvitamin D), adrenal tonus markers (serum lipids, glucose, and insulin), and renal function markers (blood urea nitrogen, creatinine, and urine analysis results).

All newborns will undergo extended hearing screening. Evoked otoacoustic emission and brainstem auditory-evoked potential tests will be performed during the child’s first year of life.

Statistical Analysis

All data will be stored in REDCap (Research Electronic Data Capture; Vanderbilt University), which is a tool for building and managing web-based surveys and databases. All variables will be summarized via standard descriptive techniques according to their type and distribution. For the bivariate
analysis, differences in categorical variables between the exposed and unexposed groups will be verified with the Chi-square test or Fisher exact test, whereas differences in continuous variables between the groups will be assessed with the Student t test or the Mann-Whitney U test.

For dichotomous outcomes, binomial regression models, which will be adjusted based on the unbalanced and relevant background features of the comparison groups, will be used to estimate the relative risks between the exposed and nonexposed groups. Partial correlation and general linear models will be used to assess the associations between continuous outcome variables and covariates; adjustments for imbalances will be made as appropriate. A P value of <.05 will be considered significant. The control group will be composed of nonexposed mother-child dyads that meet the inclusion criteria. However, those with positive serology tests (IgG and IgM tests conducted at 14-21 days postpartum) will not be included in the control group. The groups will not be matched or paired based on age or other variables. However, any differences between the groups will be adjusted later via statistical means.

Ethics Approval and Consent to Participate

The PROUDEST was approved by the Research Ethics Committee of the University of Brasilia School of Medicine (Certificado de Apresentação de Apreciação Ética 32359620.0.0000.5558) [45]. It was also registered in the Brazilian Register of Clinical Trials [46]. All pregnant women participating in the PROUDEST are required to sign an informed consent form to join the PREGNANT branch. Likewise, the participation of the children in the BORN branch will require signed, informed consent from their mothers. The 6-month reports on the study’s status and its partial results will be made available to the institutional Research Ethics Committee and may be publicly consulted upon request.

Availability of Data and Materials

At the time of the publication of this protocol, study enrollment and data collection have already started, but we have not completed the participant recruitment and data analysis phases. Therefore, data sharing is not yet feasible, as no data sets have been generated or analyzed at this stage of the study. As partial data become available, they will be displayed in the Brazilian Register of Clinical Trials [46].

Results

The PROUDEST is in the data collection phase. Study recruitment started in July 2020. As of January 2021, a total of 260 pregnant women who were infected with SARS-CoV-2 during pregnancy and 180 newborns from hospitals in the Federal District in Brazil have been included in the study. Data analysis is scheduled to start after all data are collected.

Discussion

Study Implications

The PROUDEST offers comprehensive insight (from both the obstetric and pediatric perspectives) into the effects of SARS-CoV-2 infection on the global health of pregnant women and their offspring. Specifically, the study will fill the deep gap in knowledge about the consequences of SARS-CoV-2 infection during early gestation (ie, the period when the critical stages of embryogenesis take place), as women in all stages of pregnancy will be followed. The virus-induced inflammatory and immunological phenomena that occur in SARS-CoV-2–exposed mothers during this early period of life may have a particular impact on placental and fetal physiology or may even be associated with epigenetic signals. Therefore, these phenomena could conceivably affect the long-term outcomes of a child’s growth, development, and metabolism.

A better understanding of these potential long-term consequences requires lengthy, prospective observational studies, such as the PROUDEST. This study will not only address the clinical outcomes associated with in-pregnancy exposure to COVID-19 but also evaluate a host of soluble tissue biomarkers (as described in this protocol) with the aim of comprehensively understanding the mechanisms underlying related clinical phenomena.

Our data will add to the overall clinical and basic knowledge base for COVID-19, and our ultimate goal is to provide grounds for better managing SARS-CoV-2–exposed pregnant women and their children through direct means or by setting the stage for additional studies. In fact, the PROUDEST opens up a broad spectrum of possibilities for further, multidisciplinary research on the effects of SARS-CoV-2 infection on maternal, fetal, and pediatric health. Furthermore, our results might prove to be relevant from a social perspective, as they may provide data that support the tailored development and implementation of health policies that are specifically oriented to this particular demographic group.

The study does have several limitations. Its observational nature limits inferences for causal associations to some extent. However, the cohort study design is the closest observational equivalent to a clinical trial in terms of analytical power, and the objective of the PROUDEST does not ethically allow for interventional experiments because such experiments would imply that pregnant woman will be randomized based on SARS-CoV-2 infection. Moreover, the purpose of the study is to characterize the clinical and pathophysiological phenomena associated with SARS-CoV-2 infection during pregnancy and infancy, not to test the efficacy of any intervention. Thus, we believe that a cohort study is the best possible study design for addressing our objectives. The lack of random allocation for comparison groups will be partially compensated by statistically adjusting for the observed imbalances. Vaccination for SARS-CoV-2 will not be an exclusion criterion because due to the vaccine’s limited accessibility and availability, there are no feasible methods for estimating the proportion of pregnant women that will receive the vaccine. However, the effects of SARS-CoV-2 vaccination can be adjusted and analyzed in small control groups that have either received or not received the vaccine.

The protracted follow-up in the BORN substudy is expected to result in the dropout of several participants. We set a sample size for the study after taking into consideration a 20% loss to follow-up rate. This might ensure that a sizable number of
children are available for the final assessment when they reach age of 5 years. However, we cannot avoid survival bias in the long-term data. Nevertheless, given that the continuity of multidisciplinary assistance will be guaranteed for all children in the BORN branch throughout the study period regardless of their withdrawal from the analysis or (temporary) losses to follow-up, we expect that children experiencing health problems associated with in-pregnancy COVID-19 exposure will be less likely to drop out than those who are in perfect health. Therefore, we do expect to have a sufficient number of children in the long run for identifying developmental abnormalities (should they exist), even after some amount of dropout.

Several routine pediatric consultations will be emphasized in the study, such as checking the kind of alimentation that a child is receiving (human milk or formula milk) and promoting the practice of breastfeeding.

Conclusions
The PROUDEST is a long-term, prospective cohort study designed to provide a comprehensive analysis of the effects of in-pregnancy exposure to COVID-19 on women and their offspring from a clinical and pathophysiological standpoint. Our results might contribute to the improvement of the management of SARS-CoV-2-exposed mother-child dyads—through direct means or by setting the stage for future related studies—by providing knowledge on the clinical-pathophysiological phenomena associated with COVID-19 exposure among this particular population.

Acknowledgments
We are immensely grateful for the contributions of all participants among the mother-child dyads, medical students, residents, and health professional staff from the participating hospitals; their support is essential for conducting our study. The resources used in outpatients’ follow-ups, ambulatorial consultations, and complementary exams (diagnostic imaging and laboratorial studies) are provided by the University Hospital of Brasilia, Asa Norte Hospital of Brasilia, and the Central Laboratory of Public Health of the Federal District. All public institutes are funded by the Brazilian health care system (Sistema Único de Saúde).

Authors’ Contributions
GF, LS, and FM are cofirst authors. GMF, LS, FM, ACZ, and LMHM drafted and finalized the manuscript. CPA, GM, LS, ACZ, and MECC provided input. AASS, COA, DAAJ, JALJ, RMT, LCGC, and CPA critically reviewed the manuscript. All authors made insightful contributions. All authors have read and approved the manuscript.

Conflicts of Interest
None declared.

References


Abbreviations

CCL: CC motif chemokine ligand
CNS: central nervous system
CSF: cerebrospinal fluid
CXCL: CXC motif chemokine ligand
IgG: immunoglobulin G
IgM: immunoglobulin M
IL: interleukin
MERS-CoV: Middle East respiratory syndrome coronavirus
PROUDEST: Pregnancy Outcomes and Child Development Effects of SARS-CoV-2 Infection Study
REDCap: Research Electronic Data Capture
RT-PCR: reverse transcriptase–polymerase chain reaction
SARS-CoV: severe acute respiratory syndrome coronavirus
Digital Health Intervention to Increase Health Knowledge Related to Diseases of High Public Health Concern in Iringa, Tanzania: Protocol for a Mixed Methods Study

Christine Holst¹, MPhil; Felix Sukums², PhD; Bernard Ngowi³,⁴, MD, PhD; Lien My Diep⁵, MSc; Tewodros Aragie Kebede⁶, MSc; Josef Noll⁷,⁸, PhD; Andrea Sylvia Winkler⁹,¹⁰, MD, PhD

¹Centre for Global Health, Department of Community Medicine and Global Health, Institute of Health and Society, University of Oslo, Oslo, Norway
²Directorate of ICT, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania
³Muhimbili Medical Research Centre, National Institute for Medical Research, Dar es Salaam, United Republic of Tanzania
⁴Mbeya College of Health and Allied Sciences, University of Dar es Salaam, Mbeya, United Republic of Tanzania
⁵Oslo Centre for Biostatistics & Epidemiology, Oslo University Hospital, Oslo, Norway
⁶Fafo Institute for Labour and Social Research, Oslo, Norway
⁷Basic Internet Foundation, Kjeller, Norway
⁸Department of Technology Systems, University of Oslo, Oslo, Norway
⁹Centre for Global Health, Institute of Health and Society, University of Oslo, Oslo, Norway
¹⁰Center for Global Health, Department of Neurology, Technical University of Munich, Munich, Germany

Corresponding Author:
Christine Holst, MPhil
Centre for Global Health, Department of Community Medicine and Global Health
Institute of Health and Society
University of Oslo
Postboks 1130 Blindern
Oslo, 0318
Norway
Phone: 47 48234044
Email: christine.holst@medisin.uio.no

Abstract

Background: Traditionally, health promotion and health education have been provided to communities in the global south in the form of leaflets or orally by health care workers. Digital health interventions (DHIs) such as digital health messages accessed by smartphones have the potential to reach more people at a lower cost and to contribute to strengthening of health care systems. The DHI in this study focuses on disseminating digital health education regarding 3 disease complexes of high public health concern: HIV/AIDS, tuberculosis, and Taenia solium (neuro)cysticercosis or taeniasis, a parasitic zoonotic disease that requires a One Health approach. The DHI presents the participants with animated health videos (animations) and provides access to information spots (InfoSpots) with a free-of-charge digital health platform containing messages about health to rural Tanzanian communities.

Objective: The objective of this study is to measure the effect of the DHI on health knowledge uptake and retention over time in the rural communities.

Methods: This is a mixed methods study including a nonrandomized controlled trial and qualitative interviews conducted in rural Tanzania. A health platform containing digital health messages for the communities was developed prior to the study. The health messages consist of text, pictures, quizzes, and animations of everyday stories, aimed at disease prevention and early treatment. The baseline and immediate postintervention assessments were completed in Iringa, Tanzania in May 2019. The participants were interviewed by enumerators and completed questionnaires regarding health knowledge. Participants in the intervention group were exposed to 3 different health animations once on a tablet device. The participants’ health knowledge was assessed again immediately after the exposure. The first follow-up survey was undertaken in August 2019. The InfoSpots with the digital health platform were thereafter launched in the intervention villages in November 2019. Qualitative interviews were undertaken in February 2020. The second follow-up was completed in June 2020.
Results: A total of 600 participants have been enrolled in the trial. We will assess (1) the difference in knowledge scores between baseline and the immediate postintervention assessments in the intervention group and (2) the difference in knowledge scores between the intervention and control groups at baseline, 3 and 6 months post-DHI rollout. Since a randomized design did not prove feasible, potential confounders (eg, age, gender, education, and time of exposure) may be introduced, and results will be adjusted. Data analysis for the 35 qualitative interviews is currently ongoing, and perspectives and experiences related to use and nonuse of the InfoSpots are being explored.

Conclusions: The data have been collected, and the analysis is ongoing in this digital health study, aimed at evaluating the effects of a DHI based on relevant health messages. The publications of results can be expected this year.

Trial Registration: ClinicalTrials.gov NCT03808597; https://clinicaltrials.gov/ct2/show/NCT03808597

International Registered Report Identifier (IRRID): RR1-10.2196/25128

(JMIR Res Protoc 2021;10(4):e25128) doi: 10.2196/25128

KEYWORDS
digital health; eHealth; mHealth; DigI; Tanzania; digital health messages; digital health promotion; digital health education; HIV/AIDS; tuberculosis

Introduction

Background

Health education is seen as an essential component in efforts to prevent diseases [1]; furthermore, it is the part of health care that is concerned with promotion of healthy behavior [2]. Health education and health promotion are two terms that are sometimes used interchangeably [3]. Health promotion has a wider meaning, defined by the Ottawa charter in 1986 as “the process of enabling people to increase control over, and to improve, their health” [4]. The World Health Organization defines health literacy as something that “represents the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health” [5]. Health literacy is a key determinant of health; limited health literacy is an underestimated problem and is associated with risky behavior, poorer health, less self-management, and more hospitalization and health care costs [6]. Health literacy is often perceived as the outcome of health education [7].

Increased health literacy levels have been demonstrated, among others, after basic education among adults [8], tailored rehabilitation education among patients who underwent a breast operation [9], education through active learning among elders in rural areas [10], and health belief model–based health education among patients with hypertension [11]. However, additional studies are needed to provide evidence on the best context-specific approach to improve impact on health literacy, health outcomes, and ultimately health disparities.

Strategies to increase understanding and individual health literacy can include delivering only a few health messages at a time, using plain and jargon-free language, resorting to visual aids, and using the teach-back method for health care workers when talking to patients [12].

Creating awareness of the prevalence, causes and transmission, signs and symptoms, and possibilities of treatment and prevention of a disease aims to provide early treatment and increase appropriate use of health care services. This is important in order to reduce the burden of infectious diseases like HIV/AIDS, tuberculosis (TB), and Taenia solium (neuro)cysticercosis or taeniasis (TSCT) in rural areas in the global south, where access to health care services may be starkly limited. Traditionally, health education has been the task of health workers or community health workers in such areas.

Over two-thirds (25.7 million) of all people living with HIV live in regions of Africa [13]. Young women between the ages of 15 and 24 years are twice as likely to be living with HIV as men, and in sub-Saharan Africa, girls account for 5 in 6 new infections among adolescents aged 15 to 19 years [14]. Complete and reliable information, rapid diagnosis, and early treatment and care are essential to control the virus, alleviate disease symptoms, and prevent transmission. TB treatment saved over 60 million lives globally between 2000 and 2020. However, despite TB being both treatable and curable, 1.4 million people died from the disease in 2019 [15]. The most efficient way to save lives, apart from prevention, is to diagnose and treat TB as early as possible. TSCT is an emerging but neglected parasitic and zoonotic disease (ie, both humans and animals are infected) caused by the tapeworm T. solium. In humans, the disease manifests mainly as neurocysticercosis (ie, the settlement of tapeworm larvae in the brain), the most frequent preventable cause of epilepsy in T. solium endemic areas of sub-Saharan Africa [16]. TSCT, endemic in Tanzania [17], has been listed as a priority disease for international attention towards elimination and is part of a wider aim to alleviate poverty in the same regions [18]. Iringa in southern Tanzania is endemic also for the other 2 diseases covered in this study, with one of the highest HIV prevalence estimates in the country in 2016 [19] and a notification rate for TB above the national average in 2018 [20].

Today’s education channels are changing. Therefore, health education interventions must meet the target population at respective levels of access and technology [21] and rely on the characteristics of the technology available [22]. Digital health interventions (DHI) representing “a discrete functionality of the digital technology to achieve health sector objectives” [23] and electronic health (eHealth), defined as the use of information and communication technologies for health [24], are now understood as pivotal in order to provide better care to a larger...
number of people, especially those most in need [25]. eHealth can be regarded as a tool to provide the public with health messages and support as well as to motivate clients to care about health. According to the World Health Organization [23], clients are “members of the public that are potential or current users of health services, including promotion activities,” the latter playing a vital role in their health management and disease prevention decisions. eHealth includes mobile health, defined as the use of mobile wireless technologies for health [26], a useful tool to deliver education and promote health-seeking behavior and health-related lifestyle decisions, since people nowadays can be contacted more easily through their phones [27].

Within the Sustainable Development Goals endorsed by the United Nations in 2015, SGD 3, which is about ensuring healthy lives and promoting well-being for all at all ages [28], will not be achieved without giving people access to the information they need in order to live healthier lives. The use of digital tools and technology is also emphasized as a strategy for improving health literacy [29]. However, access to digital information is not enough; in order to utilize digital health resources, a basic digital literacy skill set is needed [30].

The use of smart devices and mobile data services is taken up rapidly, and it is likely that a larger part of the population in, for example, sub-Saharan Africa will have access to both of these in the near future [31]. The Tanzania Communications Regulatory Authority regularly produces reports of mobile service penetration including internet services. The reports do not provide distribution of the services by region or district. The smartphone penetration in Iringa is unknown, but in 2016, 25.3% of national mobile connections were smartphones [32]. The cost of smartphones is a critical barrier to mobile internet adoption in sub-Saharan Africa, especially among people living in rural areas [33], leaving reasons to believe that less than a quarter of the population in rural Iringa have access to other health education than via health care workers at dispensaries with various opening hours.

The Non-discriminating Access for Digital Inclusion project (hereby referred to as the DigI project) addressed this gap by providing access to free digital health messages [34] related to HIV, TB, and TSCT in 2 rural villages in Iringa, Tanzania. Eleven partners from 8 countries [35] are collaborating in the project, including local and international researchers with extensive experience within the aforementioned diseases. The multi- and interdisciplinary approach used in the development and digitization of the health messages referred to in this protocol are further described in a recent paper [36]. In the DigI project, a key performance indicator framework was developed [37], and it is assumed that the digital health intervention described in this protocol will increase digital literacy in addition to health knowledge.

Objectives
The overarching objective of this study is to assess the effect of the DHI described in the following paragraphs aiming to improve health knowledge in the rural communities of Iringa district, Tanzania. The results will contribute to the body of evidence on the impact of digital health education and promotion in a rural setting in the global south.

The effect of the DHI will be assessed by (1) the health knowledge uptake, comparing baseline assessment (T0) scores and immediate postassessment (T1) scores of the intervention group, after exposure to 3 animations (the first part of the intervention, hereby referred to as Part 1); (2) health knowledge retention, comparing baseline assessment (T0) scores with the first 3-month follow-up (T2) scores after exposure to Part 1 and with the final follow-up (T4) scores of the intervention group, 6 months post InfoSpot rollout (the second part of the intervention, hereby referred to as Part 2); (3) a comparison between the intervention and control groups regarding the changes from T0 to T2 and T4 after participation in Part 1 and Part 2; and (4) an exploration of the perspectives and experiences of the DHI among local users and nonusers using semistructured interviews (T3), after participation in Part 1 and Part 2.

Methods
This study deploys mixed methods, including a quantitative survey in a nonrandomized controlled trial, in addition to qualitative semistructured interviews.

Location
The study is conducted in the Iringa district, selected specifically because the prevalences of all 3 diseases are high and it is a rural district in which many villages do not have access to the internet. Migoli and Izazi villages were chosen as intervention villages at an early stage by the DigI team because they are easily accessible from the highway between Iringa town and Dodoma. Kimande and Idodi were chosen as control villages because they are comparable with regards to geographic, demographic, and socioeconomic features to the intervention villages and deemed to be away far enough to keep knowledge contamination by the intervention villages to a minimum. The closest intervention village (Izazi) is located about 32 km from the closest control village (Kimande). However, the road between the 2 villages spans 124 km.

See Table 1 for detailed information on the activities undertaken at the study sites, from April 2019 to June 2020. The health knowledge, health literacy, and digital literacy questions in the questionnaires are identical in T0, T1, T2, and T4.
Table 1. Activities at the study sites at each time point.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Part 1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>T2&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Part 2&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Intervention group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group involved in assessments?</td>
<td>Yes</td>
<td>Yes (supervised)</td>
<td>Yes</td>
<td>Yes (voluntarily, unsupervised)</td>
<td>Yes</td>
</tr>
<tr>
<td>Assessments</td>
<td>40 HKQ&lt;sup&gt;h&lt;/sup&gt;, 6 HLQ&lt;sup&gt;i&lt;/sup&gt;, 5 DLQ&lt;sup&gt;j&lt;/sup&gt;</td>
<td>N/A&lt;sup&gt;k&lt;/sup&gt;</td>
<td>40 HKQ</td>
<td>40 HKQ, 6 HLQ, 5 DLQ</td>
<td>N/A</td>
</tr>
<tr>
<td>Sample size, n</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300 (minus any dropouts)</td>
<td>300 (minus any dropouts)</td>
</tr>
<tr>
<td><strong>Control group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group involved in assessments?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Assessments</td>
<td>40 HKQ, 6 HLQ, 5 DLQ</td>
<td>N/A</td>
<td>N/A</td>
<td>40 HKQ, 6 HLQ, 5 DLQ</td>
<td>N/A</td>
</tr>
<tr>
<td>Sample size, n</td>
<td>300</td>
<td>N/A</td>
<td>N/A</td>
<td>300 (minus any dropouts)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<sup>a</sup>Baseline assessment, which includes baseline data collection in both groups.  
<sup>b</sup>Immediately after T0, the participants in the intervention group are exposed to 3 animations.  
<sup>c</sup>Postassessment immediately following Part 1.  
<sup>d</sup>T2 and T4 are the follow-up assessments.  
<sup>e</sup>Rollout of the second part of the intervention (InfoSpots).  
<sup>f</sup>Qualitative interviews.  
<sup>g</sup>HKQ: health knowledge questions.  
<sup>h</sup>HLQ: health literacy questions.  
<sup>i</sup>DLQ: digital literacy questions.  
<sup>j</sup>N/A: not applicable.  
<sup>k</sup>UIH: use of the InfoSpots and digital health platform.

Description of the Digital Health Intervention

The intervention described in this protocol is composed of 2 components or parts.

**Part 1**

In April and May 2019, the participants had supervised exposure to 3 animated health videos (hereby referred to as animations). The participants in the intervention group received the first part of the intervention immediately after the baseline questionnaire. They were exposed to 3 health animations on a 7-inch tablet. These animations contain health messages on (1) HIV/AIDS, (2) TB, and (3) TSCT in the local language (Kiswahili). The health messages explain aspects related to disease prevalence, causes and transmission, signs and symptoms, as well as treatment and prevention. The animations are between 3 minutes and 7 minutes long. They tell personal stories of ordinary people in the community. The stories address the particular disease and show how a person becomes infected, develops signs and symptoms, and can get medical care and how spreading of the disease can be prevented within the communities. The HIV/AIDS animation depicts the story of a woman, recently married, who discovers that she is infected with the virus. The story illustrates how one can get infected by a person one knows and trusts, but also how one can live a fairly normal life with the right medical treatment. The TB animation presents the story of a farmer who gets diagnosed with TB after showing symptoms of TB disease. The animation emphasizes disease prevention efforts covering individual, family, and community levels. The third animation tells the story of a grandfather who gets sick after consuming undercooked pork. The animation focuses on hygiene, hygienic animal handling, and medical treatment if TSCT symptoms occur. All animations can be found online, in both English and Kiswahili [34] (the English versions can be viewed in Multimedia Appendices 1-3). The health knowledge questions from the baseline questionnaire were immediately repeated after the participants in the intervention group were exposed to the animations. Subsequent to the last question, the enumerators explained where and when the information spots (InfoSpots) were planned to be installed and that the participants could access the health messages and internet for free in these local InfoSpots.
Part 2

In November 2019, we provided unsupervised community access to InfoSpots providing free health education in the form of text, pictures, quizzes, and animations.

The InfoSpots containing the digital health platform (Multimedia Appendix 4) and providing intervention villages with free access to digital health messages related to the aforementioned diseases were implemented in Tanzania in November 2019. The digital health platform is now accessible free-of-charge, either through people’s own devices, such as smartphones, laptop computers, or tablet devices, or through available community tablets in the InfoSpots, provided by the DigI project to offer access for people without their own devices. In Izazi and Migoli, InfoSpots are located in the village offices, at the local dispensaries, and at the local high school (only in Migoli). The InfoSpots are available for all clients in the intervention villages, not only for the participants in the study described in this protocol.

When clients access the digital health platform, they may choose an area of interest, for example TSCT. Here, they are able to read key messages in all areas: prevalence, causes and transmission, signs and symptoms, treatment and prevention. The key messages are illustrated with snapshots from the animations. Clients are able to answer quizzes, and the correct answers are revealed at the end of each quiz. In addition to text, there are animations for each of the 3 diseases, which can be streamed over a local server. Health messages related to other diseases are also available.

Study Design

Randomization of subvillages, households (HH), or participants to either the intervention or control group within the villages was not possible, as the InfoSpots are available for all people in the intervention villages and the risk of contamination (knowledge transfer) between participants in the 2 groups would have been too pronounced. This resulted in choosing a nonrandomized design. However, a random sample of participants was selected from the villages as described in the following paragraphs.

To complement the quantitative findings, semistructured interviews with participants in the intervention villages were included to explore experiences and perspectives related to the intervention.

Study Population

In Tanzania, 38 million of the 58 million people (65.5% of the population) live in rural areas [38]. Many of the communities are hard to access due to the poorly developed infrastructure. Iringa district, with a projected population of 254,032 in 2019 [39], is located in the southern Highlands of Tanzania bordering the Mbeya, Njombe, Morogoro, Dodoma, and Singida regions. About three-quarters of the adult population are able to read, and agricultural activities play the most important roles in employment in rural Iringa, generating close to the total gross domestic product [40]. In a study from 2017, 6 of 10 adults in Iringa were likely to have at least a primary school education [41]. Almost three-quarters of households in Iringa have earthen floors, and half of the houses have mud walls [40]. One-quarter of the households are female headed, mainly due to the HIV/AIDS epidemic [42].

The latest population census in Tanzania was in 2012 [43], with a total population of 5281 in Izazi, 10,937 in Migoli, 10,202 in Idodi, and 14,420 in Itunundu, the area that covers Kimande. The total populations in the study villages were much lower than depicted in the census according to the consulted village executive officers: 6233 people in Migoli and 1839 people in Izazi. In the control villages, the village executive officers reported the populations to be 3562 in Kimande and 3601 in Idodi.

Participants and Recruitment

Participants were recruited from subvillage lists during 2 phases. The early phase took place in April 2019. Lists of HHs were constructed at the village level based on the village register of inhabitants. In all 4 project villages, migration due to work opportunities is part of the population dynamics. We experienced that the initial lists, created at the village level, consisted of numerous HHs that had already moved to other places and that some of the randomly selected HHs were absent. Consequently, additional lists were created in May 2019, achieved by the subvillage chairpersons going door to door in their respective communities to complete and update the list of the subvillage HHs. The HHs with participants that were already enrolled in the study were removed from the updated subvillage lists but kept in the study, and additional HHs were selected to meet the required sample size. These 2 lists are the foundation of the selected HHs from each subvillage, according to probability proportionate to size [44]. Please see Multimedia Appendix 5 for the sampling plan containing calculations of the proportions to be selected from each subvillage and the flow diagram of participants from each village and subvillage in Figure 1.

The sampling unit in this study is the HH, randomly selected by a random number generator from the subvillage lists. We used the Kish grid [45], which is a sampling method using a preassigned table of random numbers when selecting the one participant per HH [46].

During the process, the subvillage chairpersons guided each of the 6 enumerators to each of the selected HHs in their respective subvillages. All communication between the research team and the participant was conducted in the local language. The enumerators were properly introduced by the village chairpersons, and the HHs were reassured of the trustworthiness of the research project by the chairpersons. Following a brief presentation of the study by the enumerator, the approached HH member, normally the head of the HH, sat with the enumerator and listed all HH members so that the enumerator could randomly select a participant from the HH via the Kish selection method.
The inclusion criteria are (1) aged 15-45 years; (2) permanently living in the randomly selected HH or, at least, planning to not move during the next 8 months; (3) being capable and willing to sign an informed consent form or, in the case of illiterate participants, to sign with their thumbprint.

The exclusion criteria are (1) aged >45 years or <15 years; (2) planning to be absent from the village for more than 3 weeks in the upcoming 8 months; (3) not being able or willing to give written informed consent or, in the case of illiterate participants, a thumbprint.

The participants in the intervention group were assessed for baseline health knowledge (T0) related to the intervention diseases in Migoli and Izazi (the intervention villages) before being exposed to the 3 animations. The latter represent the first part of the DHI, Part 1, or “description of the intervention.” Consequently, the participants of the same group were given the same questions as in the baseline questionnaire. We call this the immediate post assessment (T1). T0, Part 1 of the DHI, and T1 all took place on the same day for the each of the participants in the intervention group. The control group consists of another 300 individuals in the control villages Idodi and Kimande. In the control villages, the participants were assessed for their baseline health knowledge (T0) in the same way as in the intervention group and were given the same questionnaire.

The participants in both groups were followed up in T2 with the same questions as in T0. When the data collection from T2 was completed, village InfoSpots with access to the digital health platform with the health messages were launched in the intervention villages. This represents the second part of the DHI: Part 2. Follow-up surveys with the same questions as in T0 were repeated in both groups 6 months after the rollout of Part 2 (T4), in June 2020.

The participants in the qualitative component were randomly selected from the list of participants enrolled in the quantitative study or were recruited on site as users of the InfoSpots.

**Outcomes**

The intervention period lasted from April 2019 until June 2020. Primary outcomes of this study are focusing on health knowledge retention in the intervention group compared to the control group, 6 months after Part 2, which is 12 months after Part 1. Secondary outcomes are focusing on immediate health knowledge uptake in the intervention group only, as a result of the exposure to the animations in Part 1.

**Ethics and Informed Consent**

This protocol is registered in Helseforsk, University of Oslo, Norway, and the project has been assessed by the Norwegian Centre for Research Data (NSD), reference number 59643.
Ethical approval from the National Institute for Medical Research in Tanzania has been granted twice with the reference number NIMR/HQ/R.8a/Vol 1 IX/2947. The study is registered in ClinicalTrials.gov with ID: NCT03808597. Additionally, the first author was provided with a research permit from the Tanzanian Commission for Science and Technology.

All participants were briefed about the aims of the study, their voluntary participation, and their rights as participants according to The European Union General Data Protection Regulation (GDPR). The participants were informed about the types of personal data collected (name, sex, age, education, religion, mobile number), where the data was going to be kept, and for how long and who would have access to these data. In addition to this oral briefing, participants were given a letter containing the information, along with the contact information of the local research entity. Informed consent forms were signed by all participants willing to participate in the study and by one enumerator. The forms were developed with support from NSD.

We are fully committed to good ethical and legal practice. All information about the participants will be dealt with confidentially and will remain deidentified for 2 years after the end of the entire project. The deidentification code and the collected sensitive data will be stored separately in the University of Oslo’s Services for Sensitive Data (TSD) [47]. In 2022, the deidentification code will be deleted, and the material will remain anonymous for public reuse. A possible benefit for the participants taking part in this study is an increase in knowledge of the highly prevalent diseases of HIV/AIDS, TB, and TSCT. The InfoSpots with the digital health platform will continue to be available after the intervention period is over. No intervention was given to the control group, but when data collection ended in June 2020, the InfoSpots with the same digital health platform were implemented in the control villages.

This study focuses solely on the participants’ health knowledge, not on the participants’ own health status or personal health data. The study participants were not given any compensation for taking part in the study.

Data Collection

The data collection consists of a quantitative part, spanning from April 2019 to June 2020, and a qualitative part, conducted in February 2020.

Collection of Quantitative Data Through Questionnaires

The questionnaire (Multimedia Appendix 6) used in T0, T1, T2, and T4 contains health knowledge questions concerning all 3 diseases, covering the domains of prevalence, causes and transmission, signs and symptoms, as well as treatment and prevention. The prevalence questions explored whether the participants had heard about the diseases. In the causes and transmission domain, we questioned the participants about how the disease transmits or what the cause of the infection is. The participants were asked to describe signs and symptoms for the relevant disease and how the diseases are diagnosed. Further, we asked questions related to treatment and cure. Lastly, the participant answered questions regarding different prevention strategies, both for individuals and communities. The questions can have one or several correct answers. The enumerators read the questions to the participants but did not inform people about the possible answers. All questions are thus open-ended. The questionnaire contains questions accompanied by a list of possible answers, so the enumerator can check the answer provided by the participant.

We used 13 questions in the HIV/AIDS section: 1 related to prevalence, 4 related to causes and transmission, 2 related to signs and symptoms, 3 related to treatment, and 3 related to prevention. We also asked the participants where they could find sufficient and reliable health information on HIV/AIDS in their communities. In the next section, the participants answered 10 questions related to TB: 1 in the prevalence domain, 2 in the causes and transmission domain, 2 related to the signs and symptoms, 3 related to treatment and, finally, 2 regarding prevention. The last section was comprised of 17 questions related to TSCT. Since this is a relatively unknown disease, we came up with 4 questions related to the prevalence of the disease, in the case of both pork tapeworm and cysticercosis. We used 4 questions in the causes and transmission domain, 4 questions in the signs and symptom domain, 1 question related to treatment, and 4 questions concerning prevention. The participants also answered 6 health literacy questions, in addition to 5 digital literacy questions.

During T4, we asked the participants of the intervention group whether they had taken advantage of the InfoSpots and the digital health platform and which part of the platform they had accessed, before repeating the questionnaire.

The study questionnaire, containing health knowledge questions regarding our 3 diseases of interest, digital and health literacy, and the use of the InfoSpots, was developed within the Digl group over a 6-month period in 2018. We used validated questions from earlier studies [48] and reports (TACAIDS/Ministry of Home Affairs, 2012; International Organisation for Migration, 2014) from various collaborations with Digl team members, as much as possible.

The Digl group had developed the key messages we wanted to convey in the health animations [36] from reliable and approved health education sources. Some of the questions were derived from this work. One key message within the HIV section, for example that girls and young women are the most vulnerable group, was emphasized in the HIV animation, and the question derived from this key message simply reads, “Which group do you think is most vulnerable to HIV infection?”

When collecting data from the questionnaire, an Android App called KoBoCollect [49] was used. The questionnaire was uploaded to KoBoToolbox [50], software for conducting field data collection in challenging environments. This is a free, open-source tool that allows users to create and deploy questionnaires to be used in an area with no internet connection. Accessing KoBoToolbox is only possible with a username and password. The questionnaire was shared with the enumerators, who were able to submit data but could not access the collected information. Once the tablets were connected to Wi-Fi or mobile internet, the app automatically transferred the collected data to the online platform KoBoToolbox.
During the first survey round, the participants were given questions concerning their demographic information. This personal information — name, mobile phone number, and village — was registered on the HH list but not in KoBoCollect. Neither was it uploaded to KoBoToolbox, so to reduce the risk of data privacy violation. A digital object identifier (DOI) was created for each participant, and the DOI was registered in KoBoCollect along with age and sex. Thus, no personal information was uploaded to KoBoToolbox. The personal identification key was stored directly in TSD via the tablets and not transmitted via email.

The reason for asking the participants for their names and mobile phone numbers was to be able to relocate the participants in the follow-up surveys. During the follow-ups in T2 and T4, the research staff had lists with names and DOIs. When entering the data from the follow-ups, the enumerators only had to identify the participants according to the list and to use the DOI when collecting new data in KoBoCollect.

Collection of Qualitative Data Through Semistructured Interviews

In order to complement the quantitative data, 35 semistructured interviews with people from the intervention villages were conducted at T3 (ie, 3 months after Part 2). The main aim was to understand the respondents’ perspectives on the intervention [51]. Results from this data collection will provide in-depth information on how the intervention was perceived and received in the communities. It may also shed light on the possible effects of the intervention, raise opportunities and problems, and provide a different view on the use of the InfoSpots along with the digital health messages.

An interview guide (Multimedia Appendix 7) was prepared in order to steer the conversations and to make sure that no question remained unanswered or unclear. The interviews were conducted between clients in the intervention villages and the first author of this paper, in collaboration with a Tanzanian Kiswahili translator. The participants were picked randomly from the same lists as in the quantitative study and by convenience sampling in and around the InfoSpots, to include users not associated with the quantitative study. The semistructured interviews were recorded with a recording app from the University of Oslo and transcribed afterwards. The records and transcriptions were stored separately from the deidentification code, in TSD, to ensure anonymity and fulfill the data handling requirements of the GDPR. The deidentification code will be deleted 2 years after the project is completed. Afterwards, the data will remain anonymous in the NSD archives.

Training and Piloting of the Questionnaire

A 3-day training for the enumerators was held in Morogoro, Tanzania, in March 2019; 6 enumerators took part. The enumerators were familiarized with the tablets and the questionnaire installed on the tablets. They practiced the registration in TSD and learned how to provide the participant with the relevant information and contact details, in addition to obtaining informed consent. Mock interviews, using the tablets with a training version of the questionnaire, followed by submission of questionnaires via KoBoCollect, were staged.

The team went to the piloting site, Luhindo primary school, located about 40 minutes from Morogoro, on the second day. Partners from Sokoine University of Agriculture had arranged to pilot the questionnaire with a group of people working and living in the Morogoro area. Altogether, 50 submissions were entered into KoBoCollect.

On the final day of training, the Kish selection method was practiced. Registration of all participants from the pilot into TSD the day prior was performed, assuring that the personal data were stored carefully in TSD whilst their anonymous answers to the disease knowledge, digital literacy, and health literacy questions were transferred to KoBoToolbox.

All the posed questions were reviewed by the pilot participants. Seven questions were altered, and some alternative responses were added.

Prior to the training, the study tools had been translated from English to Swahili. After the training, the questionnaire was revised in the local language and translated back to English. This procedure was also applied to the consent forms and information letter.

Data Analysis Plan

Analysis of Quantitative Data

We will assess the effect of the DHI by calculating the number of correct answers and the knowledge scores in both groups at all set time points. The first calculation will be related to the baseline assessment (T0) in the intervention and control group, the immediate postassessment (T1) in the intervention group, followed by the number of correct answers and scores at 3 months (T2) post Part 1 and 6 months (T4) post Part 2, in both groups. The scores will be calculated with 1 point for each correct response and summarized for each disease and the various domains (eg, causes and transmission or signs and symptoms). The number of correct answers and scores will be applied when comparing the 2 groups and the same group at different time points. The number of correct answers and the scores will be used to follow individual participants over time, as well as to indicate the trends in the increase in health knowledge uptake and retention at community level.

A sample size of approximately 460 participants was required to detect a difference of 15% to 20% between 2 groups with and without intervention, assuming that the proportion of correct responses was 50% in the group without intervention. A 2-sided significance level of .016 and 80% power were used. The significance level was set at .016 (using Bonferroni correction) to compensate for 3 time points since it was not known which time point would become relevant for the comparison between the groups. We decided to include 600 participants to increase precision of the key estimates and take dropouts into account.

The data will be analyzed in Stata/SE version 16. The McNemar test for paired data will be applied to determine marginal homogeneity of 2 dichotomous variables. The scores will be assessed against other variables such as age, education level, and gender. Summary tables with calculated averages (mean,
median, and mode) will be created. We will compare between the groups at baseline, 3 and 12 months, with and without adjusting for variables that could have influenced the respective changes. Mixed-effect models will be used to estimate the difference and monitor the alterations of variables for each question, domain score, and overall score.

**Analysis of Qualitative Data**

The data from the interviews will be analyzed in the software NVivo. A coding frame will be developed, and a content analysis method will be applied [52]. Themes of interest from the interview guide, as well as new themes found in the data, will be analyzed to contextualize the participants’ perspectives and experiences. When analyzing and summarizing the interviews, a table will be developed with excerpts from the most important findings of the qualitative research, alongside quotes from the interviews.

**Results**

The data collection was completed recently, and no results can be presented at this stage. Recruitment started April 1, 2019 and ended May 25, 2019. The 600 participants recruited have been followed over 1 year. However, 106 participants dropped out before T4, leaving 494 participants for complete follow-up analysis. The InfoSpots were launched in the intervention villages in November 2019 and in the control villages in June 2020. By enabling free-of-charge use of the InfoSpots for all community members, the approach is inclusive and not only focused on empirical findings.

The 35 qualitative interviews were undertaken in February 2020. All data are currently being analyzed, and the results are expected to be published in 2021.

**Discussion**

The primary aim of this study is to evaluate the effect of the DHI on the rural population in Iringa, Tanzania.

The quantitative part will provide a broad, “hard scientific” overview of the trends in correct answers (ie, knowledge scores) over time within and between intervention and control groups. By counting the number of correct answers before and after the intervention and comparing the knowledge scores of the respective groups, we will be able to see whether the intervention has successfully increased health knowledge in the intervention group.

As the quantitative approach only focuses on numbers, not the people themselves, the team decided to also include a qualitative component. Semistructured interviews will allow the researchers to apply an informal tone with the participants, and the findings will add meanings, opinions, values, feelings, and personal experiences to the quantitative research. The interviews can also function as arenas for discussion and may contribute to the validation of the quantitative findings.

The use of mixed methods will provide a well-founded understanding of the project results [53]. Combining the quantitative with the qualitative findings will lead to a robust set of data, not just to evaluate the DHI but also as a basis for adjusting the intervention regarding implementation of both the digital health messages and the InfoSpots with access to the digital health platform. Drawing upon findings from both research approaches to improve the intervention, the project demonstrates a participatory and community-based component that has the potential to lead to a better, context-specific impact on digital health education of local communities.

**Limitations**

This study has several limitations. The results may not be applicable to the greater rural population in Tanzania, since only a small number of villages and subvillages were included in the study. Although the study accounted for anticipated dropouts by including 100 additional participants in the study sample, this is an environment in which people shift location frequently. The rate of dropouts may be larger than anticipated in T2 and T4. People in the communities migrate as a result of the season, rainfall, drought, or harvest. A high dropout rate may introduce bias. Dropout analysis will be conducted.

Another limitation is that the participants in the intervention group may have influenced other potential participants of the same group. Data collection for T0 and T1 in the intervention villages was ongoing for 2 months. It is possible that some participants in the intervention group may have contaminated the findings by talking about the health animations they were exposed to in Part 1. Also, the research team cannot control the participants from the intervention group, who might travel to the control villages and inform recruited participants about the DHI.

**Conclusions**

We anticipate that our study will contribute to increasing knowledge of HIV/AIDS, TB, and TSCT as well as increasing overall health and digital literacy for individuals and communities in Iringa, Tanzania. By providing health messages in a digital format, the participants and their immediate surroundings will increase their health-related knowledge, which has potential to lead to an adaptation in health-seeking behavior, although this was not investigated in our study. Digitization of health information aimed at people in rural areas of sub-Saharan Africa may also contribute to strengthening of health systems, especially in low-income settings.

**Acknowledgments**

This study is included as a health research component in the Ministry of Foreign Affairs/The Norwegian Agency for Development Cooperation and The Norwegian Research Council-funded innovation project “Non-discriminating access for Digital Inclusion” (DigI project).
Ahmed Madar and Ibrahimu Mdala at the University of Oslo contributed to estimating the sample size in the first phase of this project. Flora Kajuna and Helena Ngowi assisted with the training of enumerators and the piloting of the questionnaire in Morogoro.

**Conflicts of Interest**
None declared.

---

**Multimedia Appendix 1**
“The story of tapeworms” - TSCT animation.  
*[MOV File, 78062 KB - resprot_v10i4e25128_app1.mov]*

**Multimedia Appendix 2**
*[MP4 File (MP4 Video), 58029 KB - resprot_v10i4e25128_app2.mp4]*

**Multimedia Appendix 3**
*[MP4 File (MP4 Video), 28571 KB - resprot_v10i4e25128_app3.mp4]*

**Multimedia Appendix 4**
Screenshot of the digital health platform - health information.  
*[PNG File, 622 KB - resprot_v10i4e25128_app4.png]*

**Multimedia Appendix 5**
Sampling plan.  
*[DOCX File, 15 KB - resprot_v10i4e25128_app5.docx]*

**Multimedia Appendix 6**
Study questionnaire (HIV / TB / TSCT knowledge questions used at T0/T1, T2 and T4).  
*[PDF File (Adobe PDF File), 398 KB - resprot_v10i4e25128_app6.pdf]*

**Multimedia Appendix 7**
Interview guide for qualitative study.  
*[DOCX File, 12 KB - resprot_v10i4e25128_app7.docx]*

**References**


2. Education for health: a manual on health education in primary health care. World Health Organization. 1988. URL: [https://apps.who.int/iris/handle/10665/77769](https://apps.who.int/iris/handle/10665/77769) [accessed 2021-02-17]


27. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014;7:25606 [FREE Full text] [doi: 10.3402/gha.v7.25606] [Medline: 25361730]


47. Services for sensitive data (TSD). University of Oslo. URL: https://www.uio.no/english/services/it/research/sensitive-data/about/index.html [accessed 2021-01-15]


Abbreviations

- **DHI**: digital health intervention
- **DigI**: Non-discriminating Access for Digital Inclusion
- **DOI**: digital object identifier
- **eHealth**: electronic health
- **GDPR**: The EU General Data Protection Regulation
- **HH**: Household
- **NSD**: Norwegian Centre for Research Data
- **TB**: tuberculosis
- **TSCT**: *Taenia solium* (neuro)cysticercosis or taeniasis
- **TSD**: Services for Sensitive Data (at University of Oslo)
Subclinical and Clinical Outcomes in Patients Coinfected With HIV and Chronic Hepatitis B Virus From Clinical Outpatient Centers in France: Protocol for an Ambispective, Longitudinal Cohort Study

Anders Boyd1, MPH, PhD; Lorenza N C Dezanet1, MD, PhD; Raisha Kassime1, MSc; Patrick Mialihes2, MD; Caroline Lascoux-Combe3, MD; Julie Chas4, MD; Pierre-Marie Girard1,5, MD, PhD; Joël Gozlan5,6, MD, PhD; Fabien Zoulim7, MD, PhD; Constance Delaugerre8, MD, PhD; Hayette Rougier9, MSc; Karine Lacombe1,5, MD, PhD

1Sorbonne Université, INSERM, Institut Pierre Louis d’Épidémiologie et de Santé Publique, Paris, France
2Service de Maladies Infectieuses et Tropicales, Hôpital de la Croix-Rousse, Hospices Civils de Lyon, Lyon, France
3Service de Maladies Infectieuses, Hôpital Saint-Louis, APHP, Paris, France
4Service de Maladies Infectieuses, Hôpital Tenon, APHP, Paris, France
5Service de Maladies Infectieuses, Hôpital Saint-Antoine, APHP, Paris, France
6Centre de Recherche Saint-Antoine, Paris, France
7Centre de Recherche sur le Cancer de Lyon, Unité 1052, INSERM, UMR 5286, CNRS, Lyon, France
8Laboratoire de Virologie, Hôpital Saint-Louis, APHP; Université de Paris, INSERM U944, Institut de Recherche Saint-Louis, Paris, France
9Institut de Médecine et d’Épidémiologie Appliquée (IMEA), Paris, France

Corresponding Author:
Anders Boyd, MPH, PhD
Sorbonne Université, INSERM
Institut Pierre Louis d’Épidémiologie et de Santé Publique
Hôpital Saint-Antoine, SMIT
184 Rue du Faubourg St. Antoine
Paris, 75012
France
Phone: 33 1 49 28 21 49
Email: anders.boyd@iplesp.upmc.fr

Abstract

Background: Previous large-scale studies have examined the effect of chronic hepatitis B virus (HBV) infection on overall and cause-specific mortality in individuals with HIV. However, few studies have collected data on the subclinical indicators of HBV that lead to these severe outcomes in the coinfected population.

Objective: In this study, we aim to describe the procedures of a cohort study extension aimed at assessing HBV-DNA replication, serological markers of HBV (hepatitis B e antigen [HBeAg] and hepatitis B surface antigen), and liver fibrosis and how these subclinical outcomes relate to mortality in predominately tenofovir-treated, coinfected patients with HIV-HBV. We assessed the characteristics at cohort inclusion of those who participated in the cohort extension, as well as those who did not participate due to being lost to follow-up or death.

Methods: Patients with HIV and chronic HBV who completed follow-up in a prospective cohort study conducted in 4 outpatient centers (Paris and Lyon, France; 2002-2011) were invited to participate in a cross-sectional visit from November 2016 to March 2018, during which a comprehensive evaluation of HIV- and HBV-related disease was undertaken. Virological and clinical data since the previous study visit were retrospectively collected.

Results: Of the 308 individuals enrolled in the cohort, 147 (47.7%) participated in the cross-sectional visit. At this visit, most participants were HBeAg negative (111/134, 82.8% with available data), had undetectable HBV DNA (124/132, 93.9% with available data), and were undergoing antiretroviral therapy containing tenofovir disoproxil fumarate or tenofovir alafenamide (114/147, 77.6%). There were no significant differences in characteristics at cohort inclusion between those who did and did not complete the cross-sectional visit, except for a lower proportion with an AIDS-defining illness (30/147, 20.5% vs 49/161, 30.4%).
respectively; \( P = .04 \). Of the 161 nonparticipating individuals, 42 (26.1\%) died, 41 (25.4\%) were lost to follow-up and known to be alive, and 78 (48.4\%) were lost to follow-up with unknown vital status. Most differences in characteristics at cohort inclusion were observed between deceased individuals and those participating in the cross-sectional visit or those lost to follow-up. With this extension, the median follow-up time of the overall cohort is presently 9.2 years (IQR 3.4-14.6).

**Conclusions:** Extended follow-up of the French HIV-HBV cohort will provide important long-term data on the subclinical trajectory of HBV disease in the coinfected population. The biases due to the relatively high rate of those lost to follow-up need to be assessed in future studies of this cohort.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/24731

**KEYWORDS**
liver cirrhosis; hepatitis B; viral load; longitudinal studies; immunosuppression

### Introduction

#### Background

With the advent of potent antiretroviral therapy (ART), the incidence of death related to AIDS has substantially declined among individuals with HIV, leading to new causes of mortality [1]. Liver-related death has been topmost among these recent causes, due in part to chronic hepatitis B virus (HBV) co-infection, and it has been a major concern over the past decade [2].

Because high levels of plasma HBV replication are associated with liver cirrhosis, hepatocellular carcinoma, liver-related death, and overall mortality [3], the suppression of circulating virus is the primary goal of therapy. The potent anti-HBV agent tenofovir (TFV), which is able to extensively suppress HBV replication and be easily incorporated in most ART regimens, has been available since 2002 [4]. Nevertheless, there are conflicting data regarding whether the risk of severe liver-related morbidity (ie, hepatic decompensation, portal hypertension, end-stage liver disease, and hepatocellular carcinoma) and mortality are indeed reduced in coinfected versus mono-infected patients despite the use of this agent [5-10].

To characterize the risk of severe liver-related morbidity and mortality over time in the coinfected population, an adequate evaluation of HBV-DNA viral replication, HBV serological battery, and liver fibrosis is needed. These parameters are important subclinical indicators of chronic HBV disease. Unfortunately, many hospital-based cohorts, including individuals with HIV, do not have regularly collected HBV virological and serological data, either due to structural or cost issues [11-13]. Furthermore, most HIV-HBV cohort studies presently include either retrospective data or small patient sizes, wherein less potent anti-HBV agents, such as lamivudine, were administered [14-16]. Therefore, there is a need for cohort studies to collect more exhaustive data on subclinical indicators of HBV in the HIV-HBV coinfected population, especially during TFV-containing ART.

It was in this context that our research group initiated the French HIV-HBV cohort, which followed coinfected patients between 2002 and 2011 [17]. We recently invited participants of the cohort to return for a study visit in 2017-2018 with the purpose of (1) performing a cross-sectional, comprehensive evaluation of viral hepatitis and liver fibrosis and (2) retrieving biological and clinical data since their previous study visit in the cohort. The primary research aim of the cohort is to assess HBV-DNA replication, serological markers of HBV (hepatitis B e antigen [HBeAg] and hepatitis B surface antigen [HBsAg]), and liver fibrosis and assess how all 3 of these outcomes relate to mortality in predominately TFV-treated, coinfected patients with HIV-HBV.

**Objectives**

In this paper, we present a protocol of this cohort. First, we aimed to expand on the procedures of the extension of this cohort study. The individual characteristics of the overall cohort were described and compared between those who did and did not participate in the cross-sectional visit mentioned earlier. Second, to provide more accurate information on mortality, we described the procedures by which we obtained the vital status of individuals who were lost to follow-up and assessed the number of cohort participants lost to follow-up. The characteristics of individuals who continued follow-up, who were lost to follow-up, and who were deceased were then compared. Finally, we discussed the importance of these data and how they will be used in future research.

### Methods

#### Study Design and Visits

The French HIV-HBV cohort is a closed, prospective, longitudinal cohort study that included patients initially from 7 academic hospital centers located in Paris and Lyon, France, and is aimed at evaluating the determinants of liver-related morbidity and mortality in individuals coinfected with HIV-HBV. The inclusion criteria were HIV-positive serology confirmed by western blot, HBsAg-positive serology for at least 6 months, a Karnofsky score ≥70, aged 18 years or older, and provided signed written informed consent. Exclusion criteria were acute HBV infection (ie, HBsAg-positive for less than 6 months); any severe physical, clinical, or mental condition preventing participation (as determined by a Karnofsky score ≤70 or from recommendations by the treating physician); not fulfilling all inclusion criteria; or refusal to participate. Patient recruitment and follow-up of the cohort occurred in 3 phases, with the first 2 phases previously described in detail [17]. Table 1 shows the general characteristics of each study phase.

---

https://www.researchprotocols.org/2021/4/e24731
Protocols for each study phase were approved by the hospital research, consistent with the French Public Health Law. In accordance with the European AIDS Clinical Society guidelines for the clinical management and treatment of chronic hepatitis B and C coinfection in adults with HIV [18].

In the first phase, 308 patients were recruited between May 2002 and May 2003 and prospectively followed every 3 months until the month-36 visit (2005-2006). In the second phase, 185 patients completing follow-up approximately 12 to 24 months after the first phase were recruited between March 2007 and March 2008. Patients with METAVIR F3/F4 fibrosis continued study visits every 6 months and patients with METAVIR F0/F1/F2 continued study visits every 12 months until the month-36 visit (2010-2011).

In the third phase, individuals completing follow-up in either the first or second phase were invited from November 2016 to March 2018 to participate in a cross-sectional study visit, which included a comprehensive evaluation of viral hepatitis and liver fibrosis. Approximately 4 mL of additional serum and plasma samples were collected and stored at −20 °C in a centralized location (Tumérothèque, Hôpital Saint-Antoine, Paris, France). Data were also retrospectively collected from the most recent study visit of the first or second phase until the cross-sectional visit or, in the case of death, until the most recent visit before death. The retrospective visits were selected from a 3-month period before or after the date of available HBV-DNA viral load (VL) measurement or, if missing, from a 3-month period before or after the date of the yearly visit.

All patients provided written informed consent at the beginning of each study phase. Retrospective data from deceased individuals were collected only if the patient had signed consent for the use of their clinical data during noninterventional research, consistent with the French Public Health Law. Protocols for each study phase were approved by the hospital ethics committee (first phase: Hôpital Pitié-Salpêtrière; second phase: Hôpital Saint-Antoine; third phase: Hôpital Hôtel-Dieu) in Paris, France, in accordance with the Helsinki Declaration.

### Primary Outcome Measures

A total of 4 primary outcomes were studied using data from this cohort extension.

#### Plasma HBV-DNA VL

This outcome was determined at inclusion and every 6 months in the first phase, every 6 to 12 months in the second phase, and in all retrospective and cross-sectional visits of the third phase. Commercially available polymerase chain reaction (PCR)–based assays were used to quantify HBV-DNA VL (COBAS AmpliPrep/COBAS TaqMan, detection limit: 12 international units [IU]/mL or COBAS Amplicor HBV Monitor, detection limit: 60 IU/mL; Roche Diagnostics).

We intend to study HBV DNA as continuous VLs, as undetectable HBV-DNA VLs, and as cumulative time-averaged HBV-DNA VLs [17].

#### HBeAg and HBsAg Seroclearance

This outcome was obtained from a complete HBV serological battery, which was performed at inclusion and every 12 months in the first phase, every 6 to 12 months in the second phase, and all retrospective and cross-sectional visits of the third phase. This consisted of qualitative HBsAg, HBeAg, anti–hepatitis B surface antibody (HBsAb), and anti-HBe antibody (HBeAb) detected using a commercial enzyme immunoassay (DiaSorin; Monolisa HBsAg, Bio-Rad; or Architect, Abbott Diagnostics).
We intend to study HBeAg seroclearance, defined as the transition from HBeAg-positive to HBeAg-negative status during follow-up in patients with HBeAg, and HBsAg seroclearance, defined as the transition from HBsAg-positive to HBsAg-negative status during follow-up in all included patients. We also aim to study HBeAg-seroconversion (ie, HBeAg seroclearance while acquiring anti-hepatitis B e antibody) in individuals with HBeAg and HBsAg-seroconversion (ie, HBsAg seroclearance while acquiring anti-HBsAb).

Liver Fibrosis

This outcome will be determined from several sources. First, several noninvasive biochemical scores predicting liver fibrosis levels were collected at inclusion and every 12 months in the first phase, every 6 to 12 months in the second phase, at the discretion of the treating physician in the third phase retrospective visits, and at the third phase cross-sectional visit. These scores included the FibroTest (Bio-Predictive), Fibromètre (Liver-Gastroenterology Department, CHU Angers), and Hepascore, all of which have been validated for use in the HIV-HBV coinfected population [19]. Second, measures of transient elastography (TE) were obtained at the physicians’ discretion in the first phase, every 6 to 12 months in the second phase, at the physicians’ discretion in the third phase retrospective visits, and at the third phase cross-sectional visit. TE was obtained using FibroScan (EchoSens) with either M or XL probes. Only TE measures fulfilling reliability criteria, as established by the manufacturer, were retained (ie, ≥10 valid measurements, IQR <30% of median stiffness, or ≥260% success rate) [20]. Shear wave elastography was also performed for some participants using the 2D real-time shear wave (Aixplorer, SuperSonic Imaging SA) with a 3.5 MHz convex ultrasound (SCX-6-1) and 7.5 MHz linear ultrasound (SL-10-2) probes.

We intend to study liver fibrosis as a continuous measure (in which case the analysis will only focus on one noninvasive measure) or at validated thresholds of METAVIR F3-F4 liver fibrosis [19] (in which case noninvasive measures can be combined, eg, any score indicating F3-F4 fibrosis).

All-Cause Mortality

Deaths observed during follow-up in the first and second phases, along with the underlying cause of death, were reported by the treating physician. At the third phase cross-sectional visit, treating physicians were asked to ascertain the vital status (ie, alive, deceased, or unknown) of patients completing follow-up in the first or second phases. If death occurred, they were requested to provide further information on the date and underlying cause of death.

To obtain vital status for patients with unknown status (ie, lost to follow-up), a trusted third party (Inserm U1018) was requested to link data from the French HIV-HBV cohort to a national identification registry (Répertoire national d’identification des personnes physiques, fichiers n° 779 and 801). For deceased individuals, the cause of death was then obtained by a separate trusted third party (CépiDC), linking data from the French HIV-HBV cohort to a national registry of death certificates (certificat médical) and death notifications (bulletin d’Etat civil de décès). Both registries are managed by the Institut National de la Statistique et des Études Économiques.

We intend to use all-cause mortality as an outcome of the analyses. Given the expected number of deaths, we will likely be unable to study the specific causes of death; however, these will be described in detail.

Covariables

Laboratory Measures

Other laboratory measures, as listed below, were performed on blood samples taken during the study visits.

Antibodies to hepatitis C or D virus were detected using a commercial enzyme immunoassay at inclusion and every 12 months in the first and second phases, at the physician’s discretion during the third phase retrospective visits, and at the third phase cross-sectional visit. Serum hepatitis C virus (HCV) RNA and/or hepatitis D virus (HDV) RNA were quantified if the corresponding antibody test was positive (for the first and second phase and at the third phase cross-sectional visit) at the discretion of the treating physician (for the third phase retrospective visits). HCV RNA VLs were determined using either a PCR-based assay (COBAS Amplipcr HCV Monitor v2.0, detection limit: 60 IU/mL; COBAS AmpliPrep/COBAS TaqMan HCV, detection limit: 10 IU/mL, Roche Diagnostic Systems; Abbott RealTime HCV, detection limit: 12 IU/mL, Abbott Molecular Inc) or branched-DNA technique (VERSANT HCV 3.0, detection limit: 615 IU/mL, Bayer Diagnostics). HDV-RNA VLs were quantified using a real-time quantitative PCR assay (sensitivity threshold: 1000 copies/mL [21] or 500 copies/mL [22]).

HIV-related markers of replication and immunostatus were collected at each study visit for all study phases. HIV-1 VLs were measured using either a branched-DNA technique (b-DNA Quantiplex 3.0, detection limit: 50 copies/mL, Bayer Diagnostics) or real-time PCR technique (COBAS AmpliPrep/COBAS TaqMan HIV-1 Test, detection limit: 40 copies/mL, Roche Molecular Systems). CD4+ T cell counts were quantified using standard measurements, while nadir CD4+ cell counts were obtained from patient records before inclusion in the first phase.

A complete hepatic battery was performed at each study visit for all phases and included quantification of the following using standard methods: alanine aminotransferase, aspartate aminotransferase, gamma-glutamyl transferase, alkaline phosphatase, total and conjugated bilirubin, alpha-fetoprotein, ferritin, albumin, haptoglobin, α2-macroglobulin, prothrombin time, and activated partial thromboplastin time. Hyaluronic acid was measured using an enzyme-linked protein-binding assay (Hyaluronic Acid Test Kit, Corgenix). A complete cardiovascular battery was performed at each study visit for all phases and included the following markers quantified using standard methods: apolipoprotein A1, total cholesterol, high-density lipid cholesterol, low-density lipid cholesterol, and triglycerides.
Clinical Measures

Height, weight, and systolic and diastolic blood pressure were collected at each study visit for all phases as part of routine care. Alcohol consumption was assessed at inclusion and every 12 months in the first and second phases and at the third phase cross-sectional visit. Patients were asked whether they drank alcohol and, if so, how many glasses per day, week, or month were consumed on average over the past year.

Abdominal echography was performed at the physician’s discretion during all phases. The following information was extracted from patient files: presence of ascites; presence, number, and size of nodules; presence of portal thrombosis; presence of hepatomegaly; presence of steatosis; and evidence of dysmorphic liver.

Treatment

Data on antiretroviral and anti-HBV treatments, including the specific agent, start and stop dates, doses, and reason for discontinuation, were obtained during all phases. Data on all concomitant treatment, which included the same information as antiviral treatment, were also obtained during the first and second phase; however, data on only a specific set of concomitant treatments (given by Anatomical Therapeutic Chemical [ATC] Classification codes in Multimedia Appendix 1) were obtained during the third phase retrospective visits and at the third phase cross-sectional visit. All treatments were classified using the ATC codes [23]. Treatment information was retrieved from patient files and verified by the treating physician.

Clinical Events

Data on clinical events, including date of presentation, date of resolution (if applicable), and necessary intervention were obtained during all phases. Data on all clinical events were collected during the first and second phase; however, data on only a specific set of clinical events (given by International Classification of Diseases [ICD]-10 codes in Multimedia Appendix 2) were collected during the third phase retrospective visits and at the third phase cross-sectional visit. All clinical events were classified using the ICD-10 codes [24]. Information on clinical events was retrieved from patient files and verified by the treating physician.

Data Management

Data were collected using paper clinical report forms, which were filled out by trained clinical research associates. Data for the third phase retrospective and third phase cross-sectional visits were manually entered in a centralized database maintained at INSERM UMRS_1136, Hôpital Saint-Antoine.

To ensure data validity, we used a 10% random sample of study participants and double-entered their data. For any data field with >5% discordance between data entries, the entire data field was reverified, and discrepancies were resolved with the data manager.

Statistical Analysis

We described the characteristics of the study population at cohort inclusion using counts and proportions or medians and IQRs. We then compared characteristics at cohort inclusion between those participating versus not participating in the third phase cross-sectional visit using a Kruskal-Wallis test for continuous variables and a Pearson χ² test or Fisher exact test for categorical variables.

After combining vital status from study centers and external data sources, we categorized cohort participants as follows: (1) those completing the third phase cross-sectional visit, (2) discontinuation of follow-up due to death (as determined during follow-up or by linked data from CépiDC), (3) lost to follow-up with known alive vital status (as determined by clinical outpatient records obtained from the treating physician or by linked data from CépiDC), or (4) lost to follow-up with unknown vital status. We then compared participant characteristics at cohort inclusion between groups using the same statistical tests as described above.

Statistical analysis was performed using Stata software (version 15.0), and significance was determined using a P value <.05.

Results

Recruitment Flow of Participants in the Third Phase Cross-sectional Study Visit

Of the 308 patients included in the first study phase, 142 (46.1%) were no longer followed at the participating center due to death (35/142, 24.6%), lost to follow-up (69/142, 48.5%), followed by a nonparticipating center (36/142, 25.3%), incarcerated (1/142, 0.7%), or unknown reasons (1/142, 0.7%). Of the 166 patients still in care at the participating center, 5 (3.5%) were not proposed to participate in the third phase cross-sectional visit, and 13 (7.8%) refused participation. The remaining 148 patients provided informed consent to participate in the third phase cross-sectional study visit. One patient continued care in a nonparticipating center after signing consent and did not complete their visit, so 147 patients were considered to have completed the third phase cross-sectional visit. Patient flow from the beginning of the cohort inclusion until the third phase cross-sectional visit is summarized in Figure 1.
Figure 1. Patient flow. Patient numbers are given between the first and third phases of the French HIV–hepatitis B virus cohort. Reasons for study discontinuation between these phases are also provided. Seven individuals initially considered as lost to follow-up were in fact deceased based on date linked to the CépiDS database.

Participants of the Third Phase Cross-sectional Study Visit and Their Characteristics

Of the 147 patients participating, most were male (119/147, 80.9%) with a median age of 55 years (IQR 49-59) at the time of the third phase cross-sectional visit. The median time since the first positive HIV and HBsAg tests was 24 years (IQR 18-28) and 22 years (IQR 17-26), respectively. Immunosuppression was for the most part mild with median CD4+ cell count at 548/mm$^3$ (IQR 426-719), and HIV RNA was predominately undetectable (136/139, 97.8% with available data). Most patients were HBeAg negative (111/134, 82.8% with available data) and many had undetectable HBV DNA (124/132, 93.9% with available data) with a median HBV-DNA VL of $2.53 \log_{10}$ IU/mL (IQR 2.53-2.95) for those with detectable HBV DNA. Most patients (114/147, 77.6%) underwent either TFV or TFV alafenamide–containing ART.

When comparing patients who did and did not participate in the third phase cross-sectional visit (Table 2), the only difference in inclusion characteristics observed was a lower proportion with an AIDS-defining illness (30/147, 20.4% vs 49/161, 30.4%, respectively; $P=.04$).
Table 2. Characteristic of the study population at cohort inclusion, stratified by participation in the third phase cross-sectional study visit.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Phase 1 (N=308)</th>
<th>Phase 3 Participatingb (n=147)</th>
<th>Nonparticipating (n=161)</th>
<th>P valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male, n (%)</td>
<td>259 (84.1)</td>
<td>119 (81.0)</td>
<td>140 (87.0)</td>
<td>.15</td>
</tr>
<tr>
<td>Age (years), median (IQR)</td>
<td>40 (35-45)</td>
<td>40 (35-44)</td>
<td>39 (35-45)</td>
<td>.93</td>
</tr>
<tr>
<td>Alcohol consumption &gt;1 glass/day (n=295), n (%)</td>
<td>176 (59.7)</td>
<td>87 (61.2)</td>
<td>89 (58.2)</td>
<td>.59</td>
</tr>
<tr>
<td><strong>BMI (kg/m²; n=291), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.94</td>
</tr>
<tr>
<td>Underweight (16.5-18.5)</td>
<td>16 (5.5)</td>
<td>8 (5.8)</td>
<td>8 (5.3)</td>
<td></td>
</tr>
<tr>
<td>Normal (18.5-25.0)</td>
<td>221 (75.9)</td>
<td>105 (75.5)</td>
<td>116 (76.3)</td>
<td></td>
</tr>
<tr>
<td>Overweight (25-30)</td>
<td>46 (15.8)</td>
<td>23 (16.6)</td>
<td>23 (15.1)</td>
<td></td>
</tr>
<tr>
<td>Moderate or severe obesity (&gt;30)</td>
<td>8 (2.8)</td>
<td>3 (2.2)</td>
<td>5 (3.3)</td>
<td></td>
</tr>
<tr>
<td>Estimated HIV infection duration (years), median (IQR)</td>
<td>9.9 (3.6-14.0)</td>
<td>9.4 (3.8-13.1)</td>
<td>10.2 (3.6-14.6)</td>
<td>.29</td>
</tr>
<tr>
<td>AIDS-defining illness, n (%)</td>
<td>79 (25.6)</td>
<td>30 (20.4)</td>
<td>49 (30.4)</td>
<td>.04</td>
</tr>
<tr>
<td>CD4+ T cell count (per mm³), median (IQR)</td>
<td>400 (269-555)</td>
<td>404 (283-557)</td>
<td>399 (264-554)</td>
<td>.83</td>
</tr>
<tr>
<td>Nadir CD4+ T cell count (per mm³; n=271), median (IQR)</td>
<td>212 (103-325)</td>
<td>212 (107-309)</td>
<td>215 (90-344)</td>
<td>.81</td>
</tr>
<tr>
<td>ARTc experienced at inclusion, n (%)</td>
<td>251 (81.8)</td>
<td>122 (83.0)</td>
<td>129 (80.6)</td>
<td>.59</td>
</tr>
<tr>
<td>Detectable HIV RNA (&gt;50 copies/mL), n (%)</td>
<td>145 (47.2)</td>
<td>64 (43.5)</td>
<td>81 (50.6)</td>
<td>.21</td>
</tr>
<tr>
<td>HIV-RNA viral load (log₁₀ copies/mL)</td>
<td>3.90 (2.59-4.44)</td>
<td>3.91 (2.48-4.41)</td>
<td>3.89 (2.87-4.50)</td>
<td>.68</td>
</tr>
<tr>
<td>From country of high endemicity, n (%)</td>
<td>86 (27.9)</td>
<td>47 (32.0)</td>
<td>39 (24.2)</td>
<td>.16</td>
</tr>
<tr>
<td>Estimated HBV⁵ infection duration (years), median (IQR)</td>
<td>6.1 (2.2-10.8)</td>
<td>7.3 (2.8-11.2)</td>
<td>5.3 (2.2-10.4)</td>
<td>.14</td>
</tr>
<tr>
<td>HBV DNA &gt;60 IU/mL, n (%)</td>
<td>238 (77.2)</td>
<td>113 (76.9)</td>
<td>124 (77.5)</td>
<td>.90</td>
</tr>
<tr>
<td>HBV DNA viral load (log₁₀ IU/mL)</td>
<td>4.26 (2.55-6.58)</td>
<td>3.39 (2.36-6.58)</td>
<td>4.68 (2.66-6.69)</td>
<td>.12</td>
</tr>
<tr>
<td>HBeAgf positive, n (%)</td>
<td>160 (52.0)</td>
<td>73 (49.7)</td>
<td>87 (54.0)</td>
<td>.44</td>
</tr>
<tr>
<td><strong>HBV genotype ( n =170), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.18</td>
</tr>
<tr>
<td>A</td>
<td>105 (61.8)</td>
<td>51 (68.0)</td>
<td>54 (56.8)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1 (0.6)</td>
<td>0 (0.0)</td>
<td>1 (1.1)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>17 (10.0)</td>
<td>7 (9.3)</td>
<td>10 (10.5)</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>19 (11.2)</td>
<td>8 (10.7)</td>
<td>11 (11.6)</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>19 (11.2)</td>
<td>5 (6.7)</td>
<td>14 (14.7)</td>
<td></td>
</tr>
<tr>
<td>Mixed A/D or A/G</td>
<td>9 (5.3)</td>
<td>4 (5.3)</td>
<td>5 (5.3)</td>
<td></td>
</tr>
<tr>
<td>Precore W28stop mutation (n=164), n (%)</td>
<td>47 (28.7)</td>
<td>18 (24.7)</td>
<td>29 (31.9)</td>
<td>.31</td>
</tr>
<tr>
<td>Lamivudine-resistance mutations (n=146), n (%)</td>
<td>97 (66.4)</td>
<td>42 (64.6)</td>
<td>55 (67.9)</td>
<td>.67</td>
</tr>
<tr>
<td><strong>Other viral hepatitis, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td>.48</td>
</tr>
<tr>
<td>Anti-HCVg positive serology</td>
<td>19 (6.2)</td>
<td>9 (6.1)</td>
<td>10 (6.2)</td>
<td></td>
</tr>
<tr>
<td>Anti-HDVh positive serology</td>
<td>12 (3.9)</td>
<td>6 (4.1)</td>
<td>6 (3.7)</td>
<td></td>
</tr>
<tr>
<td>Anti-HCV and anti-HDV positive serology</td>
<td>12 (3.9)</td>
<td>3 (2.0)</td>
<td>9 (5.6)</td>
<td></td>
</tr>
</tbody>
</table>

aSignificance was determined using the Kruskal-Wallis test for continuous variables and Pearson χ² test or Fisher exact test for categorical variables.
bParticipating patients were defined as those signing written informed consent and completing the third phase cross-sectional study visit. All characteristics reported in the table are from data collected at the inclusion visit of the French HIV-HBV cohort (2002-2003).
cART: antiretroviral therapy.
dAmong patients with detectable viral loads.
eHBV: hepatitis B virus.
Lost to Follow-up and Deceased Participants of the Cohort Study and Their Characteristics

Of the 161 patients not followed up at the third phase cross-sectional visit, we were able to establish vital status for the 55 patients who were already known to have been deceased (n=35), incarcerated (n=1), not proposed or refused to participate (n=5 and 13, respectively), or did not complete the third phase study visit (n=1). The remaining 106 patients not followed at the third phase cross-sectional visit did not have known vital status because of an unknown reason for lost to follow-up (n=69), being followed at a nonparticipating center (n=36), or unknown reason for noninclusion (n=1). Of these, 28 patients were able to be linked to the CépiDC database: 21 were still known to be alive and 7 were identified as deceased from the death certificate or notification registry. These deaths occurred between 2007 and 2015.

From a total of 308 participants, 147 (47.7%) participants completed the third phase cross-sectional visit, 42 (13.6%) discontinued follow-up due to death, 41 (13.3%) were lost to follow-up with known alive vital status, and 78 (25.3%) were lost to follow-up with unknown vital status. When comparing characteristics at the inclusion visit of the cohort, there were several significant differences across these patient groups (Table 3). Most differences were observed in deceased individuals who, respectively compared with all others, were significantly more likely to be older (median age 44 vs 40 years; \(P<.001\)), were not from a country of high HBV endemicity (4/42, 10% vs 82/266, 30.8%; \(P=.003\)), have a longer duration since first HIV-positive serology (median 13.1 vs 9.4 years; \(P<.001\)), have a higher proportion with AIDS-defining illness (21/42, 50% vs 58/266, 21.8%; \(P<.001\)), and have a lower nadir CD4+ cell count (median 128 vs 221 per mm\(^3\); \(P=.03\)). No significant differences in characteristics at cohort inclusion were observed between individuals completing the third phase cross-sectional visit versus lost to follow-up with known vital status. There were also no significant differences in characteristics at cohort inclusion between individuals completing the third phase cross-sectional visit versus lost to follow-up with unknown vital status, with the exception of longer duration since first HBsAg-positive serology (median 7.3 vs 3.9 years, respectively; \(P=.009\)) and higher proportion with detectable HIV RNA (45/78, 58% vs 64/147, 43.5%, respectively; \(P=.04\)).
Table 3. Characteristics of the study population at cohort inclusion, stratified by those completing the third phase cross-sectional visit, deceased, and lost to follow-up with known or unknown vital status.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Completed follow-up (n=147)</th>
<th>Deceased&lt;sup&gt;a&lt;/sup&gt; (n=42)</th>
<th>Lost to follow-up (n=119)</th>
<th>Known vital status&lt;sup&gt;+&lt;/sup&gt; (n=41)</th>
<th>Unknown vital status (n=78)</th>
<th>P value&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male, n (%)</td>
<td>119 (81.0)</td>
<td>39 (92.9)</td>
<td>35 (85.4)</td>
<td>66 (84.6)</td>
<td></td>
<td>.32</td>
</tr>
<tr>
<td>Age (years), median (IQR)</td>
<td>40 (35-44)</td>
<td>44 (38-53)</td>
<td>39 (34-41)</td>
<td>38 (34-43)</td>
<td></td>
<td>.002</td>
</tr>
<tr>
<td>Alcohol consumption &gt;1 glass per day (n=295), n (%)</td>
<td>87 (61.3)</td>
<td>22 (55.0)</td>
<td>22 (55.0)</td>
<td>45 (61.6)</td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td><strong>BMI (kg/m&lt;sup&gt;2&lt;/sup&gt;; n=291), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.99</td>
</tr>
<tr>
<td>Underweight (16.5-18.5)</td>
<td>8 (5.8)</td>
<td>3 (7.5)</td>
<td>1 (2.5)</td>
<td>4 (5.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal (18.5-25.0)</td>
<td>105 (75.5)</td>
<td>31 (77.5)</td>
<td>31 (77.5)</td>
<td>54 (75.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight (25-30)</td>
<td>23 (16.6)</td>
<td>5 (12.5)</td>
<td>7 (17.5)</td>
<td>11 (15.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate or severe obesity (&gt;30)</td>
<td>3 (2.2)</td>
<td>1 (2.5)</td>
<td>1 (2.5)</td>
<td>3 (4.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated HIV infection duration (years), median (IQR)</td>
<td>9.4 (3.8-13.1)</td>
<td>13.1 (7.1-15.6)</td>
<td>10.1 (2.8-14.6)</td>
<td>9.4 (3.2-13.8)</td>
<td></td>
<td>.06</td>
</tr>
<tr>
<td>AIDS-defining illness, n (%)</td>
<td>30 (20.4)</td>
<td>21 (50.0)</td>
<td>12 (29.3)</td>
<td>16 (20.5)</td>
<td></td>
<td>.001</td>
</tr>
<tr>
<td>CD4+ T cell count (per mm&lt;sup&gt;3&lt;/sup&gt;, median (IQR))</td>
<td>404 (283-557)</td>
<td>370 (249-474)</td>
<td>411 (338-619)</td>
<td>406 (268-562)</td>
<td></td>
<td>.65</td>
</tr>
<tr>
<td>Nadir CD4+ T cell count (per mm&lt;sup&gt;3&lt;/sup&gt;; n=271), median (IQR)</td>
<td>212 (107-309)</td>
<td>128 (65-304)</td>
<td>235 (172-346)</td>
<td>238 (70-384)</td>
<td></td>
<td>.08</td>
</tr>
<tr>
<td>ART&lt;sup&gt;+&lt;/sup&gt; experienced at inclusion, n (%)</td>
<td>122 (83.0)</td>
<td>37 (88.1)</td>
<td>31 (75.6)</td>
<td>61 (79.2)</td>
<td></td>
<td>.45</td>
</tr>
<tr>
<td>Detectable HIV RNA (&gt;50 copies/mL), n (%)</td>
<td>64 (43.5)</td>
<td>14 (34.2)</td>
<td>22 (53.7)</td>
<td>45 (57.7)</td>
<td></td>
<td>.05</td>
</tr>
<tr>
<td>HIV-RNA viral load (log&lt;sub&gt;10&lt;/sub&gt; copies/mL), median (IQR)</td>
<td>3.91 (2.48-4.41)</td>
<td>3.83 (2.59-4.05)</td>
<td>3.99 (3.04-4.53)</td>
<td>3.67 (2.87-4.46)</td>
<td></td>
<td>.89</td>
</tr>
<tr>
<td>From country of high endemicity, n (%)</td>
<td>47 (32.0)</td>
<td>4 (9.5)</td>
<td>13 (31.7)</td>
<td>22 (28.2)</td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>Estimated HBV&lt;sup&gt;+&lt;/sup&gt; infection duration (years), median (IQR)</td>
<td>7.3 (2.8-11.2)</td>
<td>6.7 (2.3-13.5)</td>
<td>5.5 (2.8-11.5)</td>
<td>3.9 (1.7-8.9)</td>
<td></td>
<td>.04</td>
</tr>
<tr>
<td>HBV DNA &gt;60 IU/mL, n (%)</td>
<td>113 (76.9)</td>
<td>31 (75.6)</td>
<td>31 (75.6)</td>
<td>62 (79.5)</td>
<td></td>
<td>.95</td>
</tr>
<tr>
<td>HBV-DNA viral load (log&lt;sub&gt;10&lt;/sub&gt; IU/mL), median (IQR)</td>
<td>3.39 (2.36-6.58)</td>
<td>4.63 (2.71-6.88)</td>
<td>5.02 (2.28-6.58)</td>
<td>4.54 (2.76-6.88)</td>
<td></td>
<td>.45</td>
</tr>
<tr>
<td>HBeAg&lt;sup&gt;+&lt;/sup&gt; positive, n (%)</td>
<td>73 (49.7)</td>
<td>29 (69.0)</td>
<td>20 (48.8)</td>
<td>38 (48.7)</td>
<td></td>
<td>.13</td>
</tr>
<tr>
<td><strong>HBV genotype (n=170), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.36</td>
</tr>
<tr>
<td>A</td>
<td>51 (68.0)</td>
<td>17 (60.7)</td>
<td>13 (61.9)</td>
<td>24 (52.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1 (2.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>7 (9.3)</td>
<td>4 (14.3)</td>
<td>2 (9.5)</td>
<td>4 (8.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>8 (10.7)</td>
<td>1 (3.6)</td>
<td>1 (4.8)</td>
<td>9 (19.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>5 (6.7)</td>
<td>4 (14.3)</td>
<td>4 (19.1)</td>
<td>6 (13.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed A/D or A/G</td>
<td>4 (5.3)</td>
<td>2 (7.1)</td>
<td>1 (4.8)</td>
<td>2 (4.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precore W28stop mutation (n=164), n (%)</td>
<td>18 (24.7)</td>
<td>4 (15.4)</td>
<td>7 (33.3)</td>
<td>18 (40.9)</td>
<td></td>
<td>.10</td>
</tr>
<tr>
<td>Lamivudine-resistance mutations (n=146), n (%)</td>
<td>42 (64.6)</td>
<td>14 (66.7)</td>
<td>14 (77.8)</td>
<td>27 (64.3)</td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td><strong>Other viral hepatitis, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.31</td>
</tr>
<tr>
<td>Anti-HCV&lt;sup&gt;+&lt;/sup&gt; positive serology</td>
<td>9 (6.1)</td>
<td>2 (4.8)</td>
<td>4 (9.8)</td>
<td>4 (5.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-HDV&lt;sup&gt;+&lt;/sup&gt; positive serology</td>
<td>6 (4.1)</td>
<td>1 (2.4)</td>
<td>3 (7.3)</td>
<td>2 (2.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Specifically coinfected patients with HIV-HBV may hamper its toxicity issues related to treatment and/or comorbidities among shown effective antiviral potency [4]. However, long-term coinfection patients [17]. TFV-containing ART regimens have [29] and is achieved in approximately 1% per year in treated follow-up important. HBsAg seroclearance is a slow process for up to 15 years of follow-up. There are certain facets of HIV-HBV coinfection that make data sets containing longer end points to ensure sufficiently powered comparisons. Third, despite the large number of patients and extensive follow-up, we might not have enough end points to ensure sufficiently powered comparisons. Third, certain data were not consistently collected during follow-up (eg, TE measurements only began at the start of the second phase). As a result, any future analysis using available samples will have a gap in follow-up time.

**Discussion**

**Principal Findings**

A handful of prospective cohort studies have examined the effect of HIV-HBV coinfection on mortality outcomes [10], but most of these studies have incomplete data on HBV status, making it difficult to determine how HBV infection contributes to HBV-related disease. We present the design of a recent extension of a large observational cohort of patients coinfected with HIV-HBV in France to address the shortcomings of these previous cohorts.

**Strengths and Weaknesses**

One of the major strengths of our cohort is the consistent collection of HBV-DNA VLs, complete HBV serological battery, HBV genetic characterization (specifically regarding antiviral resistance mutations in the pol gene, vaccine escape mutations in the S gene, and W28* precore mutations [25]), and liver fibrosis measurements. To our knowledge, few cohorts have regularly collected these data and those that have this information are no longer in active follow-up [26,27]. The regular collection of markers of HBV replication and HBV serological data will help provide more precise estimates of when certain events occur (eg, time to undetectable HBV-DNA or HBeAg or HBsAg seroclearance) as well as their durability (eg, persistent HBV-DNA viremia or sustained HBeAg or HBsAg seroclearance) or variability (as was observed for noninvasive measures of liver fibrosis [28]). At the same time, we aim to associate these primary outcomes with all-cause mortality.

In addition, by extending our previous cohort, we collected data for up to 15 years of follow-up. There are certain facets of HIV-HBV coinfection that make data sets containing longer follow-up important. HBsAg seroclearance is a slow process [29] and is achieved in approximately 1% per year in treated coinfected patients [17]. TFV-containing ART regimens have shown effective antiviral potency [4]. However, long-term toxicity issues related to treatment and/or comorbidities among specifically coinfected patients with HIV-HBV may hamper its clinical effectiveness. It is also unknown how HIV-HBV coinfection evolves in an aging population, especially with respect to severe non-AIDS and non-HBV-related clinical events. The longer follow-up provided in the French HIV-HBV cohort could help establish the frequency of these end points and their determinants, provided that there are a sufficient number of events.

Another advantage of our cohort is the availability of frozen samples dating back to the beginning of the cohort. As these samples have been stored at a centralized location, they have facilitated their use in several collaborations, some of which have involved the genetic variability of HBV [25,30], compared with patients undergoing intensified with pegylated-interferon [31], HDV-RNA replication in tri-infected patients [32], kinetics of HBsAg or HBeAg quantification [33], hepatitis B core-related antigen quantification [34], and covalently closed circular DNA [35] during TFV treatment and ongoing projects related to markers of replication during treatment. Nevertheless, it should be mentioned that stored samples are only available in the first and second phases and at the third phase cross-sectional visit. As a result, any future analysis using available samples will have a gap in follow-up time.

There are other limitations in the French HIV-HBV cohort. First, most patients had years of ART experience before inclusion in the first phase, were HBeAg positive, came from predominately Western Europe, and sexually acquired HBV. The vast majority of individuals with HIV currently diagnosed with an HBV coinfection in Europe come from regions of high HBV endemicity, while new HBV infections have declined dramatically in other Western European settings [36,37]. Therefore, our study population may not be fully representative of modern coinfected patients. Second, despite the large number of patients and extensive follow-up, we might not have enough end points to ensure sufficiently powered comparisons. Third, certain data were not consistently collected during follow-up (eg, TE measurements only began at the start of the second phase) or across the entire study population (eg, HBsAg levels are only available in the first 2 phases among TFV-treated
patients). Fourth, metabolic data (insulin, adiponectin, controlled attenuation parameter of the FibroScan, etc) were not collected at any point during follow-up, making it difficult to understand the increasing role of metabolic disorders in this patient population [38].

Assessment of Participation in the Third Phase Study Visit and Loss to Follow-up

Similar to other longitudinal studies, one major concern is loss to follow-up and the possible biases induced thereto. At the third phase of follow-up, occurring roughly 15 years after the start of the cohort, almost half of the patients did not continue follow-up. When compared with rates of loss to follow-up in HIV-positive (5.4% of patients lost to follow-up from 1999 to 2009 in the French Hospital Database on HIV study [39]) or HIV-HCV coinfected (10% of patients lost to follow-up from 2006 to 2010 in the Agence Nationale de Recherche sur le Sida et les Hépatites (ANRS) CO13 HEPAVIH study [40]) cohorts from France, loss to follow-up would seem to occur at a much higher rate in our cohort.

Individuals lost to follow-up could have a higher risk of more advanced disease (opportunistic infections, advanced liver fibrosis or cirrhosis, cancer, advanced cardiovascular diseases, etc), which might make them less willing to participate or continue follow-up. No particular population characteristic at cohort inclusion was significantly different between participating and nonparticipating patients across phases, except for a higher proportion of AIDS-defining illnesses in those not participating in the third phase cross-sectional visit. There could also be other characteristics during follow-up, either measured or unmeasured (eg, nonreported comorbidities) in our cohort that were associated with loss to follow-up. This might induce differential loss to follow-up bias, the effect of which depends on the research question addressed. Therefore, the assessment of this bias is imperative in future studies from our cohort.

Previous studies linking data from the French Hospital Database on HIV study and other mortality outcome registries or cohorts demonstrated that 29.8% of patients with HIV who were lost to follow-up had in fact died [41]. In our cohort, this proportion was similar at 32% (7/22), and notwithstanding the fact that 61.9% (78/126) of individuals with lost to follow-up were unable to have their vital status obtained is comparable with the proportion deceased of those remaining in follow-up (35/182, 19.2% either died or participated in the third phase cross-sectional visit).

Conclusions

Long-term and more comprehensive cohorts of coinfected individuals with HIV-HBV are pivotal in understanding how liver-related and nonliver-related morbidity and mortality can develop in this patient population. We were able to extend our cohort data on HBV replication, HBV serology, and liver fibrosis to up to 15 years; however, given the fact that approximately half of the patients were unable to participate in the third phase cross-sectional study, either due to death or lost to follow-up, the biases associated with this high rate of loss to follow-up need to be considered in future studies from this cohort. Nevertheless, these data will be helpful in enhancing our knowledge of the clinical trajectory during coinfection. More specifically, the clinical implications of persistent viremia, lack of HBeAg or HBsAg seroclearance, and progressive development of liver fibrosis will be evaluated along with their association with all-cause mortality. The French HIV-HBV cohort may also contribute to future collaborations aimed at assessing rarer outcomes, such as hepatic decompensation, hepatocellular carcinoma, and cause-specific death.

Acknowledgments

The authors are grateful to the patients and clinical teams for their commitment to the French HIV-HBV cohort study. The authors acknowledge L Roguet for managing the logistics of the French HIV-HBV cohort; L Melila, R Kassime, and R Loisel for their help with data entry; G Pannetier and Prof F Carrat for their help in data management; and Prof P Cales for kindly providing Fibrometre scores. Finally, the authors would like to thank Prof J-F Flejou and E Roux of the Tumorothèque HUEP at Hôpital Saint-Antoine for storing samples.

This work was supported by a grant from the ANRS. Gilead Sciences, Inc, provided an unrestricted grant for the French HIV-HBV cohort in the first phase and was not involved in any part of the data collection, analysis, and writing of the paper. A doctoral fellowship was awarded to KL from the ANRS (2003-2006) and postdoctoral fellowships from the ANRS and SIDACTION were awarded to AB (2012-2014 and 2014-2017, respectively), and a postdoctoral fellowship from the ANRS was awarded to LD (2018-2020).

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Authors' Contributions

AB, LD, and RK managed data for the French HIV-HBV cohort, performed statistical analysis, and drafted the paper; PM, CL-C, and JC helped conceptualize and design the French HIV-HBV cohort, participated in patient recruitment, and provided critical input to the paper; JG and CD validated virological and serological measures and provided critical input to the paper; FZ coordinated data collection on genetic variability and gave valuable comments on the paper; HR coordinated and participated in data collection.
and helped draft parts of the paper; P-MG and KL initiated the French HIV-HBV cohort study, coordinated data collection, and drafted parts of the paper. All authors approved the final version of the paper.

Conflicts of Interest
KL has served on advisory boards and received travel grants from Gilead, Abbvie, MSD, ViiVHealthcare, and Janssen. All other authors have conflicts to declare.

Multimedia Appendix 1
Classes of concomitant treatment collected during the third phase cross-sectional study visit.

[DOCX File, 26 KB - resprot_v10i4e24731_app1.docx]

Multimedia Appendix 2
Types of clinical events collected during the third phase cross-sectional study visit.

[DOCX File, 26 KB - resprot_v10i4e24731_app2.docx]

References


Abbreviations

ANRS: Agence Nationale de Recherche sur le Sida et les Hépatites
ART: antiretroviral therapy
ATC: Anatomical Therapeutic Chemical
HBeAg: hepatitis B e antigen
HBsAb: hepatitis B surface antibody
HBsAg: hepatitis B surface antigen
HBV: hepatitis B virus
HCV: hepatitis C virus
HDV: hepatitis D virus
ICD: International Classification of Diseases
IU: international unit
PCR: polymerase chain reaction
TFV: tenofovir
TE: transient elastography
VL: viral load

©Anders Boyd, Lorenza N C Dezanet, Raisha Kassime, Patrick Miallhes, Caroline Lascaux-Combe, Julie Chas, Pierre-Marie Girard, Joël Gozlan, Fabien Zoulim, Constance Delaugerre, Hayette Rougier, Karine Lacombe. Originally published in JMIR

https://www.researchprotocols.org/2021/4/e24731

XSL-FO RenderX
Research Protocols (http://www.researchprotocols.org), 06.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Autonomic Nervous System Maturation and Emotional Coordination in Interactions of Preterm and Full-Term Infants With Their Parents: Protocol for a Multimethod Study

Christina Koumarela¹, MA; Theano Kokkinaki², PhD; Giorgos Giannakakis³, PhD; Katerina Koutra³, PhD; Eleftheria Hatzidaki⁴, PhD

¹Laboratory of Applied Psychology, Department of Psychology, University of Crete, Rethymnon, Greece
²Institute of Computer Science, Foundation of Research and Technology, Heraklion, Greece
³Department of Psychology, University of Crete, Rethymnon, Greece
⁴Department of Neonatology, Neonatal Intensive Care Unit, School of Medicine, University of Crete, Heraklion, Greece

*all authors contributed equally

Corresponding Author:
Theano Kokkinaki, PhD
Laboratory of Applied Psychology
Department of Psychology
University of Crete
Gallos
Rethymnon, 74 100
Greece
Phone: 30 2831077536
Email: kokkinaki@uoc.gr

Abstract

Background: There is limited knowledge on the physiological and behavioral pathways that may affect the developmental outcomes of preterm infants and particularly on the link between autonomic nervous system maturation and early social human behavior. Thus, this study attempts to investigate the way heart rate variability (HRV) parameters are related to emotional coordination in interactions of preterm and full-term infants with their parents in the first year of life and the possible correlation with the developmental outcomes of infants at 18 months.

Objective: The first objective is to investigate the relationship between emotional coordination and HRV in dyadic full-term infant–parent (group 1) and preterm infant–parent (group 2) interactions during the first postpartum year. The second objective is to examine the relationship of emotional coordination and HRV in groups 1 and 2 in the first postpartum year with the developmental outcomes of infants at 18 months. The third objective is to investigate the effect of maternal and paternal postnatal depression on the relation between emotional coordination and HRV in the two groups and on developmental outcomes at 18 months. The fourth objective is to examine the effect of family cohesion and coping on the relation between emotional coordination and HRV in the two groups and on developmental outcomes at 18 months.

Methods: This is an observational, naturalistic, and longitudinal study applying a mixed method design that includes the following: (1) video recordings of mother-infant and father-infant interactions at the hospital, in the neonatal period, and at home at 2, 4, 6, 9, and 12 months of the infants’ life; (2) self-report questionnaires of parents on depressive symptoms, family cohesion, and dyadic coping of stress; (3) infants’ HRV parameters in the neonatal period and at each of the above age points during and after infant-parent video recordings; and (4) assessment of toddlers’ social and cognitive development at 18 months through an observational instrument.

Results: The study protocol has been approved by the Research Ethics Committee of the University of Crete (number/date: 170/September 18, 2020). This work is supported by the Special Account for Research Funds of the University of Crete (grant number: 10792-668/08.02.2021). All mothers (with their partners) of full-term and preterm infants who give birth between March 2021 and January 2022 at the General University Hospital of Crete (northern Crete, Greece) will be invited to participate. The researcher will invite the parents of infants to participate in the study 1 to 2 days after birth. Data collection is expected to be completed by March 2023, and the first results will be published by the end of 2023.
Conclusions: Investigating the regulatory role of HRV and social reciprocity in preterm infants may have implications for both medicine and psychology.

International Registered Report Identifier (iRRID): PRR1-10.2196/28089

(JMIR Res Protoc 2021;10(4):e28089) doi:10.2196/28089

KEYWORDS
preterm infants; heart rate variability; emotional coordination; developmental outcomes

Introduction

Preterm infants are at increased risk of a range of developmental outcomes at neurological, cognitive, social competence, socioemotional, and behavioral levels [1-3]. Both behavioral and physiological pathways may be connected with the adverse developmental outcomes of preterm infants. At the behavioral level, preterm birth disrupts the emergence of parent-infant emotional synchrony [4-8]. At the physiological level, the early disruption of autonomic nervous system (ANS) development limits its capacity to respond to the environment [2]. Heartrate variability (HRV), that is, the distribution of the time interval between successive heartbeats (R-R intervals), constitutes a measure of ANS functional maturation and sympathetic nervous system (SNS)/parasympathetic nervous system (PNS) interplay, and is an index of the PNS [2,9,10]. HRV is under the control of the PNS branch of the ANS. The vagus nerve is the main nerve of the PNS. Parasympathetic activity is referred to as vagal tone and shows the contribution of the vagus nerve to cardiac functioning [9]. Gestational age in preterm infants greatly correlates with HRV parameters, with a lower gestational age associated with lower HRV [11,12]. Further, the PNS is part of the motor pathway that is connected to facial muscles that support successful interpersonal engagement, such as facial expressions [13]. Thus, higher vagal activity has been connected with better social skills, and cardiac vagal tone monitoring in young infants constitutes an index of their capacity to regulate emotional states via facial expressivity [14,15].

In addition, ex utero third trimester development in a preterm newborn is vulnerable to developmental disruption from a variety of environmental influences [2]. This is important because parents of preterm children experience higher levels of depression and anxiety, and poorer family functioning in infancy and early childhood when compared with parents of full-term children [16,17].

The relation between a young infant’s facial expressions of emotion and autonomic activity has been rarely investigated [4,15,18]. To our knowledge, only two studies have focused on the relationship between vagal activity and dyadic mother-infant regulatory processes, and they showed that (1) the cardiac vagal tone of 6-month-old infants was linked to more symmetrical features versus disruptive patterns of communication in mother-infant dyads [19] and (2) newborn vagal activity predicts mother-infant synchrony at 3 months [20].

Infant ANS maturation along with maternal mental health (depression and anxiety) seem to intervene in the relationship between HRV/vagal activity and infant or maternal emotional expressions. In a previous study [4], the vagal tone of preterm infants was shown to be lower than that of full-term infants. Maternal behavior (gaze, affect, touch, and talk) was more frequent among preterm infants with high vagal tone than those with low vagal tone, while preterm infants with low vagal tone received the lowest amount of maternal behavior at 3 months. In the same group, infant vagal tone along with maternal depressive symptoms were predictive of mother-infant gaze synchrony.

Newborns of mothers with high levels of depressive symptoms and depressed mothers had lower vagal tone than newborns of nondepressed mothers [21,22]. Further, in the course of interactions of 3-month-old infants with their nondepressed and low-depressed mothers, infant vagal tone was associated with positive affective states. Infants with greater vagal activity expressed more positive affective states. Relevant correlations were not found in the depressed group, which implies decoupling between vagal tone and facial expressivity in infants of depressed mothers [23]. Lower vagal tone was evidenced in 6-month-old infants of depressed mothers compared with nondepressed mothers, but it was not noted at 3 months. Higher vagal tone at 6 months was related to more positive vocalizations. Though a large increase in baseline vagal activity occurred from 3 to 6 months in infants of nondepressed mothers, this increase did not occur in infants of depressed mothers [24]. Further, newborns of high anxiety mothers (compared with low anxiety mothers) had lower vagal tone [25].

To our knowledge, fathers were included in only two relevant studies. First-time fathers’ heart rates increased during interactions with their newborn infants in a controlled hospital setting. Changes in fathers’ cardiovascular physiology were not related to paternal smiling [26]. Further, neonatal vagal tone predicted gaze synchrony in interactions of 3-month-old full-term and preterm infants with their mothers and fathers, despite its marginal contribution to father-infant synchrony in full-term infants. Maternal, but not paternal, postnatal depression predicted parent-infant gaze synchrony at 3 months [4].

A limited number of studies has investigated the association between parental functioning with infant emotional and physiological regulation. Six-month-old infants in families with higher marital conflict showed lower baseline vagal tone and poorer emotional regulation [27]. Further, 6-month-old infants coming from families with higher levels of parental conflict showed lower mean levels of baseline respiratory sinus arrhythmia (RSA, ie, heart rate variation according to inspiration [acceleration] and expiration [slowing down] [2,9]), a lesser degree of RSA withdrawal in response to mothers’ still face, and a lesser degree of RSA activation in contexts in which infants were interacting with mothers. No relevant effects were
found among parent conflict, infants’ behavior, mothers’ positive affect, and dyadic synchrony [28].

The proposed research, which is the first of its kind in Greece, is interdisciplinary between psychology and medicine in the basic idea, in theoretical frameworks that drive it, in the methodology, and in the implications of the results. In particular, the proposed research is interdisciplinary in the basic idea because it focuses on both expressive behaviors (facial expressions of emotion and adapted coordination) and physiological indexes of ANS (HRV). In addition, the features of preterm infants invite the psychology-medicine collaboration because (1) they possess behavioral characteristics and have neurological immaturity that may contribute in making them difficult interactive partners [29] and increase their health risks [11] and (2) there is a steady increase in neonatal survival rates as the advent of modern intensive care for preterm infants is related to a growing concern for preterm infants’ developmental outcomes and quality of life [1]. Further, this research is interdisciplinary in theoretical frameworks because it is conceived in accordance with the psychobiological theory of innate intersubjectivity developed by Trevarthen et al [30-32] and the biobehavioral polyvalent theory developed by Porges et al [13,33,34]. The psychobiological theory of innate intersubjectivity posits that infants possess innate motives expressed with qualities of emotion and adapted to perceive, respond to, attract, and influence how other persons feel and what they, in response, will perceive and do. Sensitivity for regulation of both subjective and intersubjective impulses and feelings undergoes age-related changes attributable to developments in the body and in the motivating processes of the brain. In connection to this, the polyvalent theory links the maturational shifts in neural regulations of the ANS with infant self-regulatory and social engagement skills [13]. In addition, this study is interdisciplinary in the methodology. In particular, this is a multimeth, observational, naturalistic, and longitudinal study. The methodological strategy will include (1) video recordings of the dyadic mother-infant and father-infant face-to-face interaction in the neonatal period (at hospital) and at 2, 4, 6, 9, and 12 months (at the family’s home); (2) physiological measures, that is, HRV of infants at the neonatal period and at each of the above age points; (3) one observational instrument in order to assess the toddler’s social and cognitive development at 18 months; and (4) self-report questionnaires on postnatal depression, family cohesion, and coping between partners. Finally, investigating the regulatory role of HRV/social reciprocity in preterm infants may have implications for both medicine and psychology. In particular, different neonatal pathological states have been associated with a reduction in HRV in preterm infants, and an improvement in health conditions is followed by changes in HRV, which can be used as a possible prognostic marker [11]. Further, the findings of this study may extend our understanding on the early onset of neuropsychiatric disorders given that this may be related to factors that alter ANS maturation and limbic system functions. In addition, the findings of this study may inform interventions to promote the development of preterm infants given that children at increased risk for neuropsychiatric disorders are those with a history of prematurity [2].

We will use detailed measures of timing to analyze the way emotional coordination in spontaneous interactions of preterm and full-term infants with their parents at home in Greece (across the first year of life) is related to HRV parameters and the possible correlation with the developmental outcomes of infants at 18 months. This study extends beyond the current state-of-the-art knowledge in the following ways:

1. The inclusion of microanalysis of face-to-face father-preterm infant interaction may help in the understanding of the psychological aspects of fathers of preterm infants given that fathers face different emotional experiences than mothers after preterm birth [35,36].
2. When young infants are observed responding to sensitive caregivers with mutual concern in a familiar environment, the infants show active emotional initiative with a great variety of facial expressions of emotions [37], and they have a different motivation as compared with that in a laboratory setting [38].
3. The longitudinal design of this study is important given that regulation of intersubjective feelings undergoes age-related changes attributable to developments in the body and in the motivating processes of the brain in the first postpartum year [31].

In the absence of relevant studies in Greece and given that certain cultural elements may have the potential to promote preterm infants’ development, investigating the factors that affect the development of preterm infants is important because preterm births constitute a major public health issue [39]. The main aim of this study is to investigate the way HRV parameters are related to emotional coordination in interactions of preterm and full-term infants with their parents in the first year of life and the possible correlation with the developmental outcomes of infants at 18 months.

**Methods**

**Selection Criteria**

**Inclusion Criteria**

Full-term infants (group 1) are eligible to be included in this study if they have no medical complications. Preterm infants (group 2) are eligible to be included in this study if they meet the criterion of gestational age (32-37 weeks). We will focus on preterm infants, rather than extremely preterm infants, given variations between the two groups on communicative abilities and perceived social support of mothers [40,41].

Mothers and fathers are eligible to be included in this study if they meet the following criteria: (1) both are born and have grown up in Greece; (2) they are married to each other; (3) both have given their consent to participate in the study; (4) at least one parent is employed; and (5) both are older than 20 years of age. Greek nationality of participant parents is important given that certain cultural elements may have the potential to promote preterm infants’ development, investigating the factors that affect the development of preterm infants is important because preterm births constitute a major public health issue [39]. The main aim of this study is to investigate the way HRV parameters are related to emotional coordination in interactions of preterm and full-term infants with their parents in the first year of life and the possible correlation with the developmental outcomes of infants at 18 months.

**Exclusion Criteria**

Infants will be excluded from the study if (1) they have perinatal asphyxia; (2) they have neurological pathologies; (3) they have a history of prematurity [2].
experience malformation syndromes and major congenital malformations; (4) they have sensory deficits; (5) they present metabolic genetic disease; (6) they have central nervous system infection; or (7) they have a birth weight less than 2500 g in case of full-term birth [42].

Mothers and fathers will be excluded from the study if (1) they are not of Greek nationality; (2) they have a psychiatric illness; (3) they have issues with drug or substance abuse; (4) they are not living together; (5) they are not biological parents; and (6) they are a same-sex couple. Participation in this study will presuppose that both infants and parents will meet the inclusion and exclusion criteria.

**Recruitment**

All mothers (with their partners) of full-term and preterm infants who will give birth between March 2021 and January 2022 at the General University Hospital of Crete (northern Crete, Greece) will be invited to participate in the study. The researcher will invite the parents of infants to participate in the study 1-2 days after birth.

A minimum of 50 mothers, 50 fathers, and 50 infants (N=150) will participate in the study in two groups. Group 1 will include 25 parents and their infants born at full term (≥37 weeks gestational age) with no medical complications. Group 2 will include 25 parents and their preterm infants (32-37 weeks gestational age). The two groups will be matched for demographic variables and infant gender. This sample size is in line with previous relevant studies in Greece and in other countries, and it is adequate to perform parametric statistical tests [35,43].

**Demographic and Socioeconomic Characteristics of the Sample**

Before hospital discharge, data will be collected from the mothers on the parents’ demographic and socioeconomic characteristics, mothers’ obstetric history, and infants’ characteristics (gestational age, weight, height, head circumference, and Apgar score at 1 and 5 minutes). In the case of preterm infants, information will also include the duration and the conditions of infant hospitalization (in case of hospitalization).

**Objectives, Procedure, and Measures**

The first objective is to compare the relationship between emotional coordination and HRV in dyadic full-term infant–parent (group 1) and preterm infant–parent (group 2) interactions in the course of the first postpartum year.

For the first objective of this study, each full-term neonate/infant–parent dyad will be video recorded at six age points, that is, at the neonatal period (in the hospital) and at five important age-related transitions in the development of communication in the first postpartum year [31] (2, 4, 6, 9, and 12 months [chronological age]) at home. For preterm neonates/infants, video recordings will be carried out around 36 weeks postmenstrual age at the hospital and at 2, 4, 6, 9, and 12 months (corrected age) at home [44].

In particular, before discharge, at the hospital, each neonate-parent dyad will be video recorded for a 5-minute period. Before, during, and after each infant-parent video recording, the neonate’s HRV recording will be obtained by measuring short-term variability, which provides important information about the maturation of the ANS in newborns [11]. Before the video recording, the HRV evaluation will be performed under resting conditions in the supine position for 5 minutes, and infants will be in this position at least 5 minutes before the measurement (gold standard [9]). Similarly, after the end of the video recording, we will evaluate HRV parameters for 5 minutes. Whenever possible, we will control for the time of the day of the assessment, and we will follow consistent procedures for all participants [9].

A validated portable electrocardiography (ECG) device will be used to collect ECG recordings. ECG is considered the gold standard providing accurate measurements to perform the appropriate HRV analysis offline [45].

The order of mother-infant and father-infant interactions will be counterbalanced with a 5-minute rest for the neonates between the two video recordings. Parents will be told “Play with your baby.” Video recordings at the neonatal period will be scheduled according to mothers’ and fathers’ availability and the infants’ alertness. Parent-neonate free interactions will be carried out on the third postpartum day for the full-term group and the day prior to discharge for the preterm group.

For video recordings at home, parents will be told “Play as you normally do with your baby.” The recording will take place in a room and in a position chosen by the parent, prohibiting any third-party intervention. Given that infant state and age constitute factors that influence arousal, attention, and affect [46], we will vary the duration of video recordings across the age range of this study. Thus, for younger infants aged 2 and 4 months, each video recording will last 8 minutes in order to respect their early signs of agitation after brief face-to-face exchanges with their parents [32]. In the meanwhile, we will video record interactions of parents with older infants (aged 6, 9, and 12 months) for 10 minutes in order to predict periods of infant exploration of surroundings, which intensify after the middle of the Period of Games I (3-6 months, characterized by a transformation of the infant’s motivation, which results in both increasing exploration of the environment and effective manipulation of objects) [47], and to ensure an 8-minute face-to-face interaction. The total duration for parent-infant interactions will be 4600 minutes of video data, equally distributed between periods with the mother and father and between the two groups. HRV will be evaluated following the same procedure as in the neonatal period.

The second objective is to examine the relationship of emotional coordination and HRV in groups 1 and 2 in the first postpartum year with the developmental outcomes of infants at 18 months. At 18 months, the social and cognitive development of toddlers will be assessed at home through the administration of Bayley Scales of Infant and Toddler Development, 3rd edition [48].

The third objective is to investigate the effect of maternal and paternal postnatal depression on the relation between emotional

https://www.researchprotocols.org/2021/4/e28089
coordination and HRV in the two groups and on developmental outcomes at 18 months.

According to a relevant literature review (mentioned in the Introduction), limited information on the relation between mood disorders and HRV is focused mainly on depression rather than on other mood disorders (eg, anxiety). Considering this, parents will be asked to complete (1) the Edinburgh Postnatal Depression Scale (EPDS [49]) adapted for the Greek population [50,51] (at 2 months), which is an instrument that screens women/men for postnatal depression, and (2) the Beck Depression Inventory-II (BDI-II [52]) adapted for the Greek population [53] (at 18 months), which is a self-report scale assessing symptoms of depression.

The fourth objective is to examine the effect of family cohesion and coping on the relation between emotional coordination and HRV in the two groups and on developmental outcomes at 18 months. At 2 and 18 months, parents will be asked to complete (1) the Greek version of the Family Adaptability and Cohesion Evaluation Scale (FACES-IV Package) [54-56] to measure family functioning in terms of cohesion and flexibility and (2) the Greek version of the Dyadic Coping Inventory (DCI) [57,58] for the assessment of dyadic stress coping among the couple.

**Data Analysis**

**Video Coding**

Coding of parent-newborn interactions will be carried out using the Mother-Newborn Coding System of the Coding Interactive Behavior (CIB) Manual developed by Feldman in 1998 [59]. For each 30-second period, four maternal/paternal categories and one newborn category will be coded. Categories for the mother/father include gaze, affect, talk, and touch. Infant state is coded as fussy, cry, drowsy, sleep, alert to mother/father, or alert to the environment. Maternal/paternal affiliative behavior will include the sum proportions of maternal/paternal gaze at the infant’s face, positive affect, vocalization, and affectionate touch. Emotional/social coordination will be examined according to conditional probabilities that come from the time of maternal/paternal affiliative behavior during moments when the newborn is in an alert state to the mother/father or to the environment, and the time the mother/father expresses affiliative behavior when the infant is in another state.

Microanalysis of infant and parental facial expressions of emotions at 2, 4, and 6 months will be carried out according to a coding system that has been constructed by TK [60-62]. This system adapted and extended certain parts of the categorization system developed by previous authors [63-65]. In particular, microanalysis and coding of infant-parent emotional coordination at 2, 4, and 6 months will be carried out as follows. First, parental infant-directed speech and infant (vocalizations/nonspeech) sounds will be transcribed from the video recording by the researcher. Thereafter, they will be checked for accuracy by an assistant. Second, written verbatim accounts of parental infant-directed speech will be classified into “content categories” and then into “thematic sequences” and “focus categories” (to an accuracy of 1/25th of a second). Third, each focus category and within thematic sequence will be segmented into units and subunits of analysis, according to the duration of the pause preceding and following each thematic sequence. If the pause between successive subunits/thematic sequences is shorter than or equal to 2 seconds, these will be grouped within one unit. Fourth, within each subunit of analysis, infant and maternal/paternal facial expressions of emotions will be coded according to the type of facial expressions of emotions, the shifts in the frequency of facial expressions of emotions, and the direction of intensity change. Thereafter, interpersonal engagement categories according to the types of facial expressions of emotions (happy, interest, neutral, and sad or withdrawn) will be coded as matching, completion, and nonmatching. Interpersonal engagement categories according to the shifts in the frequency of facial expressions of emotions will be coded as synchrony. Interpersonal engagement categories according to the direction of intensity change will be coded as attenuation. In matching, one partner expresses the type of facial expression of the emotion of the other partner. In completion, one partner expresses a positive valence of facial expression of emotion (“pleasure” or “interest”) in immediate response to the other. In synchrony, the two partners match the timing of change of emotional expressions with each other. In attenuation, one partner expresses the shifts in the direction of the emotional intensity of the other partner. Emotional coordination is signified by these aspects. Instances in which the two partners do not match the type of facial expression of emotion or do not synchronize their shifts of emotions, or one partner is not attune to the direction of the emotional intensity of the other partner will be termed as nonmatching. Each emotional coordination category will be indexed by two conditional probabilities, for example, for emotional matching, the infant expresses pleasure given that the mother expresses pleasure and the infant expresses pleasure given that the father expresses the same facial expression of emotion (pleasure). These probabilities provide evidence of the co-occurrence of the same facial expression of emotion between the two partners and show the proportion of time out of the entire interaction that both the parent and the infant match their emotional expressions. For microanalysis of parent-infant emotional coordination at 9 and 12 months, all necessary adjustments will be made to the above coding scheme according to changes in the infant’s motivation for communication and exploration of the environment after 6 months. Concurrent with the microanalysis of infant-parent emotional coordination, HRV parameters will be measured according to the analysis described in the subsection “HRV Analysis” below.

**Interobserver Reliability**

To measure interobserver reliability (Cohen κ), a second observer who will be trained in the use of the coding system, but will not be aware of the hypotheses under investigation, will score a random sample of 33% of the video files. Interobserver reliability will be assessed separately for the types of facial expressions, the frequency of shifts of emotions, and the categories of emotional intensity.
Quantitative Data Analysis

HRV Analysis

Heart activity will be recorded concurrently with video recordings (ie, 5 minutes, 8 minutes, and 10 minutes) with a sampling frequency of 256 Hz. The recording length is selected as it has been proven that normally HRV parameters need at least 5 minutes in order to have sufficient statistical power [66]. However, enhanced spectral resolution using autoregressive analysis will be performed. Thus, sliding temporal windows of Δt=30 s will be investigated in order to track HRV temporal dynamics. This temporal window length is sufficient for the temporal evolution of HRV as referred in relevant studies [67,68].

Artifacts (mainly due to body motion) will be automatically suppressed. Time series with less than 10% of artifacts will be included for analysis [69]. In addition, ectopic beats will be identified and will be excluded from the analysis [70]. The HRV parameters in the time domain and the frequency domain, as well as nonlinear indices will be analyzed [9,45,66]. In connection with the focus of this study on infant vagal tone, certain relevant HRV indices will be of interest for analysis. Thus, for the time domain, we will analyze mean heart rate, SDNN (standard deviation of all RR intervals), RMSSD (root mean square of successive differences), and pNN50 (percentage of adjacent RR intervals with a duration difference greater than 50 ms). The rMSSD and pNN50 indices reflect vagal tone, and SDNN reflects the SNS and PNS [9,69].

For the frequency domain, low frequency (LF, 0.04-0.15 Hz), high frequency (HF, 0.15-0.40 Hz), and the LF/HF ratio will be analyzed. HF reflects vagal tone, while LF and the LF/HF ratio reflect a mix of sympathetic and vagal activity [9,69]. Autoregressive analysis will be preferred as it provides enhanced spectral resolution compared with Fast Fourier Transform–based spectrum analysis for the frequency domain of HRV parameters [9,67,71]. Given that infants breathe faster, bands will be adjusted accordingly, and thus, we will move the boundaries of the bands to 0.24-1.04 Hz at rest [72].

For the nonlinear HRV analysis, the following parameters derived from the Poincaré plots, which are assumed to be indicators of vagal activity [9], will be estimated: SD1 (reflecting the PNS), SD2 (reflecting the SNS), and SD1/SD2 (the relationship between the PNS and SNS) [69]. Nonlinear indices are complementary HRV indicators to enhance the estimation of the participants’ physiological status [9].

Statistical Analysis

The first objective of this study is to investigate differences in emotional coordination and HRV parameters between the two groups of infants (normal and preterm) over the first year of life. For the assessment of preterm infants’ social interaction capability, mother/father-infant dyads will be analyzed in order to evaluate the “emotional coordination” and “nonmatching” of facial expressions of emotions during spontaneous interactions [62]. First, temporal statistical analysis (analysis of variance [ANOVA]) will be used to evaluate parent-infant synchrony (through HRV/emotional coordination measurements) for each group of infants across the first year [20,73]. Second, trend analysis for HRV measurements at 2, 4, 6, 9, and 12 months will be performed in order to observe patterns of developmental changes from the first month to the end of the first year. Third, correlation analysis will be conducted to examine the intercorrelations between the variables of the study [20,74]. Lastly, we will explore whether synchrony of the interaction is more prevalent among the group of infants demonstrating higher HRV on dividing the data into high and low HRV groups [20,74].

The second objective is the investigation of possible correlations between emotional coordination/HRV and developmental outcomes at 18 months of age. A correlation analysis will be used to measure potential linear correlation between emotional coordination/HRV and developmental outcomes [75]. In addition, two-way ANOVA and multivariate analysis of variance (MANOVA) will be utilized to examine the potential effect of early mother/father-infant interactive patterns and HRV measures on infant developmental outcomes at 18 months of age [1,74].

The third objective focuses on the examination of the mediating role of parental postpartum depression and family functioning/stress coping in relation to emotional coordination/HRV and developmental outcomes at 18 months. The Wilcoxon test and bivariate correlations will be used to detect potential different patterns of the study variables between normal and high-risk groups under this condition [4,28,76,77]. Analysis of covariance (ANCOVA) will be used to compute differences in the study variables between the term and preterm infant groups [4,78]. To examine the moderating effect of parental postpartum depressive symptoms, family functioning, and stress coping on emotional coordination/HRV and the developmental outcomes of toddlers, a hierarchical linear regression analysis will be performed [4,20].

Results

The timeline of the study procedure is depicted in Figure 1.

The study protocol has been approved by the Research Ethics Committee of the University of Crete (number and date of decision: 170/September 18, 2020). This work is funded by the Special Account for Research Funds of the University of Crete (grant number: 10792-668/08.02.2021).

All mothers (with their partners) of full-term and preterm infants who will give birth between March 2021 and January 2022 at the General University Hospital of Crete (northern Crete, Greece) will be invited to participate in the study. Data collection is expected to be completed by March 2023, and the first results will be published by the end of 2023.
Discussion

There is evidence that higher vagal activity is connected with better social skills, and cardiac vagal tone monitoring for young infants constitutes an index of their capacity to regulate emotional states via facial expressivity [14,15]. In the meanwhile, a limited number of studies show that infant ANS maturation along with maternal/paternal mental health (depression) and interparental functioning seem to intervene in the relationship between HRV/vagal activity and infant or maternal/paternal emotional expressions [4,21-24,27]. This may have long-term implications for infant cognitive and socioemotional development. In order to fill this gap, the main aim of this study is to investigate the way HRV parameters are related to emotional coordination in interactions of preterm and full-term infants with their parents in the course of the first postpartum year. Toward this direction, our study has four objectives. The first objective is to investigate the relationship between emotional coordination and HRV in dyadic full-term infant–parent (group 1) and preterm infant–parent (group 2) interactions in the course of the first postpartum year. The second objective is to examine the relationship of emotional coordination and HRV in groups 1 and 2 in the first postpartum year with the developmental outcomes of infants at 18 months. The third objective is to investigate the effect of maternal and paternal postnatal depression on the relation between emotional coordination and HRV in the two groups and on developmental outcomes at 18 months. The fourth objective is to examine the effect of family cohesion and coping on the relation between emotional coordination and HRV in the two groups and on developmental outcomes at 18 months.

The proposed research is interdisciplinary in the implications of the results. In particular, investigating the regulatory role of HRV and social reciprocity in preterm infants may have implications for both medicine and psychology. With regard to medicine, different neonatal pathological states have been associated with a reduction in HRV in preterm infants, and an improvement in health conditions is followed by changes in HRV, which can be used as a possible prognostic marker. With regard to psychology, the findings of this study may extend our understanding on the early onset of neuropsychiatric disorders given that this may be related to factors that alter ANS maturation and limbic system functions. Further, the results of this study may contribute to the development of neonatal telemedicine services with the aim to educate new parents and to improve the care of preterm and full-term neonates [79].

Acknowledgments

This work is supported by the Special Account for Research Funds of the University of Crete (grant number: 10792-668/08.02.2021).

Authors' Contributions

CK was involved in the rationale for the study and the general study design as a PhD student. TK serves as a research project supervisor and designed the coding system for the microanalysis of the infant-parent interaction. EH focused on the physiological
parameters of this study. KK focused on family functioning and depression measurements. GG designed the plan of statistical analysis and is involved in HRV parameter measurements. CK, TK, EH, KK, and GG wrote the manuscript.

Conflicts of Interest
None declared.

References


Abbreviations

- **ANCOVA**: analysis of covariance
- **ANOVA**: analysis of variance
- **ANS**: autonomic nervous system
- **BDI**: Beck Depression Inventory
- **CIB**: Coding Interactive Behavior
- **DCI**: Dyadic Coping Inventory
- **ECG**: electrocardiography
- **EPDS**: Edinburgh Postnatal Depression Scale
- **FACES**: Family Adaptability and Cohesion Evaluation Scale
- **HF**: high frequency
- **HRV**: heart rate variability
- **LF**: low frequency
- **MANOVA**: Multivariate Analysis of Variance
- **PNN50**: percentage of adjacent RR intervals with a duration difference greater than 50 ms
- **PNS**: parasympathetic nervous system
- **rMSSD**: root mean square of successive differences
- **RSA**: respiratory sinus arrhythmia
- **SDNN**: standard deviation of RR intervals
- **SNS**: sympathetic nervous system
Autonomic Nervous System Maturation and Emotional Coordination in Interactions of Preterm and Full-Term Infants With Their Parents: Protocol for a Multimethod Study

Please cite as:
Koumarela C, Kokkinaki T, Giannakakis G, Koutra K, Hatzidaki E

©Christina Koumarela, Theano Kokkinaki, Giorgos Giannakakis, Katerina Koutra, Eleftheria Hatzidaki. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 12.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
**Protocol**

**Pesticide Exposure of Residents Living Close to Agricultural Fields in the Netherlands: Protocol for an Observational Study**

Daniel M Figueiredo 1, Ing; Esmeralda J M Krop 1, PhD; Jan Duyzer 2, PhD; Rianda M Gerritsen-Ebben 2, MSc; Yvonne M Gooijer 3, Ir; Henk J Holterman 4, Ir, PhD; Anke Huss 1, PhD; Cor M J Jacobs 5, Ir, PhD; Carla M Kwits 6, Ir; Roel Kruijne 5, Ir; Hans J G J Mol 7, PhD; Arné Oerlemans 8, MSc; Pieter J J Sauver 9, MD, PhD; Paul T J Scheepers 8, Ir, PhD; Jan C van de Zande 4, Ir; Erik van den Berg 5, Ir, PhD; Marcel Wenneker 6, Ir; Roel C H Vermeulen 1,10, Ir, PhD

1 Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands
2 TNO Urban Environment and Safety, Utrecht, Netherlands
3 CLM Onderzoek en Advies BV, Culemborg, Netherlands
4 Wageningen Plant Research, Wageningen University & Research, Wageningen, Netherlands
5 Wageningen Environmental Research, Wageningen University & Research, Wageningen, Netherlands
6 Schutteelaar & Partners, Wageningen, Netherlands
7 Wageningen Food Safety Research, Wageningen University and Research, Wageningen, Netherlands
8 Department for Health Evidence, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands
9 Department of Pediatrics, University Medical Center Groningen, Groningen, Netherlands
10 Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht University, Utrecht, Netherlands

**Corresponding Author:**
Daniel M Figueiredo, Ing
Institute for Risk Assessment Sciences
Utrecht University
Yalelaan 2, 3rd Fl
Utrecht, 3584 CM
Netherlands
Phone: 31 621867972
Email: d.m.figueiredo@uu.nl

**Abstract**

**Background:** Application of pesticides in the vicinity of homes has caused concern regarding possible health effects in residents living nearby. However, the high spatiotemporal variation of pesticide levels and lack of knowledge regarding the contribution of exposure routes greatly complicates exposure assessment approaches.

**Objective:** The objective of this paper was to describe the study protocol of a large exposure survey in the Netherlands assessing pesticide exposure of residents living close (<250 m) to agricultural fields; to better understand possible routes of exposure; to develop an integrative exposure model for residential exposure; and to describe lessons learned.

**Methods:** We performed an observational study involving residents living in the vicinity of agricultural fields and residents living more than 500 m away from any agricultural fields (control subjects). Residential exposures were measured both during a pesticide use period after a specific application and during the nonuse period for 7 and 2 days, respectively. We collected environmental samples (outdoor and indoor air, dust, and garden and field soils) and personal samples (urine and hand wipes). We also collected data on spraying applications as well as on home characteristics, participants' demographics, and food habits via questionnaires and diaries. Environmental samples were analyzed for 46 prioritized pesticides. Urine samples were analyzed for biomarkers of a subset of 5 pesticides. Alongside the field study, and by taking spray events and environmental data into account, we developed a modeling framework to estimate environmental exposure of residents to pesticides.

**Results:** Our study was conducted between 2016 and 2019. We assessed 96 homes and 192 participants, including 7 growers and 28 control subjects. We followed 14 pesticide applications, applying 20 active ingredients. We collected 4416 samples: 1018 air, 445 dust (224 vacuumed floor, 221 doormat), 265 soil (238 garden, 27 fields), 2485 urine, 112 hand wipes, and 91 tank mixtures.

**Conclusions:** To our knowledge, this is the first study on residents’ exposure to pesticides addressing all major nondietary exposure sources and routes (air, soil, dust). Our protocol provides insights on used sampling techniques, the wealth of data...
Introduction

Background

The application of pesticides to agricultural land in the vicinity of homes has raised questions regarding health concerns from residents living nearby. Occupational pesticide exposure has been associated with different health effects including diseases of the respiratory tract [1,2], cancer [3], and neurodegenerative diseases such as Parkinson disease [4,5]. Although residents are likely exposed to lower concentrations than are occupationally exposed individuals, they are continuously exposed because of spray drift and transport of pesticides volatilizing from nearby agricultural land to their homes [6]. In addition, possible accumulation of pesticides in the home environment [7] can contribute to higher and prolonged exposure of those residents [8] compared with urban residents. In addition, in comparison with occupationally exposed workers, more vulnerable groups such as children and the elderly may be exposed in the home environment [9].

While few studies found no clear difference between outdoor air concentrations in urban and rural areas [10], several others have shown that pesticide concentrations in the air are higher close to agricultural fields [11,12] and are higher during the spraying seasons [13,14]. Both results are also true for air and dust in the indoor environment [15,16]. Additionally, when looking at internal dose (measured by biomarkers of exposure), some studies observed significant differences in pesticide exposure levels between urban and rural populations [17-19], while others did not [20,21].

Data on pesticide exposure to residents in the Netherlands are limited, even though approximately 27% of all homes are located within 250 m of at least one cultivated agricultural field as a result of the country's population density and large agricultural sector. Given the variable outcomes in the scientific literature and the lack of information on exposure levels of the Dutch (rural) population due to pesticide use on agricultural fields, the Health Council of the Netherlands advised the government to conduct research in order to fill the above-mentioned gaps of knowledge. For this, the OBO study (“Research on exposure of residents to pesticides”) was conducted.

Objectives

OBO Study

The OBO study aimed to assess the pesticide exposure of residents living close (<250 m) to agricultural fields and to better understand possible routes of environmental exposure. Since most spraying in the Netherlands is done with a downward spraying technique [22-25] and flower bulb cultivation is known to involve a large amount of pesticides [26], the focus was on pesticide exposure among residents living in the vicinity of flower bulb fields. The emphasis of the OBO study was on the assessment of residential pesticide exposure, not on potential adverse health or toxicological effects.

This Protocol

To address the above-mentioned objectives, 3 research questions were formulated:

- What are the concentrations of pesticides in the environment of residents living close to agricultural cultivation of flower bulbs compared with those living further away?
- What is the personal exposure to pesticides of residents living close to agricultural cultivation of flower bulbs compared with those living further away?
- What are the sources and routes of exposure contributing to environmental and personal exposure to pesticides in areas of flower bulb cultivation?

In this paper, we describe the OBO study, providing an outline of the methodology used to address the above-mentioned questions. We provide relevant information, as well as lessons learned, for other researchers planning to set up similar study designs, apply similar methods, and explore collaborations (eg, make use of the collected data in pooled analysis).

Study Contributions

To the best of our knowledge, this is the first time that a study in the field of residential exposure to pesticides was set up that (1) followed various spraying applications, (2) collected both environmental and personal samples, (3) targeted a wide range of pesticides (ie, insecticides, herbicides, and fungicides), and (4) was performed in different time periods (when pesticides were applied and when pesticides were not applied), allowing the comprehensive study of both spatial and temporal variations in residential pesticide exposure. Contributions of the study to the knowledge base are presented in Textbox 1.
Textbox 1. Contributions of the study to the knowledge base.

- It produced a FAIR (Findable, Accessible, Interoperable, and Reusable) data set that includes concentrations of many different pesticides in relevant matrices, such as air, dust, soil, and urine. The data set also contains detailed information collected on spraying applications (ie, frequency, mixture applied, quantity, etc.). It can be used in future studies for multiple aims, for example to determine the more common pesticide mixtures in the environment or, together with data from other studies, develop robust models to estimate the concentrations of pesticides in certain matrices (eg, indoor home dust as a result of the take-home route).
- It adds to the growing knowledge of pesticide distribution in the environmental matrices and its main determinants, not only for sprayed pesticides but also for some pesticides that were not reported to have been applied.
- It provides valuable and useful information for other researchers and biological monitoring studies regarding toxicokinetics of some pesticides in the human body.
- It provides insights into associations between different matrices (ie, relative pesticide content of air, dust, urine, etc). These results add to the scientific evidence by bringing new knowledge to light.
- A modeling framework was developed that comprises verified models that explain the most relevant pesticide fate processes (eg, spray drift and evaporation), as well as exposure routes (eg, dermal and inhalation). Verification is possible by comparing measured values in the different matrices with modeled values, using all of the collected information on spraying applications and meteorological conditions. This framework or parts of it can be used in future studies as an exposure assessment tool.
- Spray drift and volatilization experiments were performed to increase understanding of the abovementioned processes. These experiments emphasized the importance of drift-reducing nozzles as an exposure reduction factor and the importance of volatilization in pesticide release from the fields.
- It addressed some important current knowledge gaps regarding exposure of residents. Examples are the relative pesticide concentrations in outdoor air and indoor air and which exposure routes contribute most to personal exposure. In our modeling framework, we compared the 4 main exposure routes: contact with surfaces, dust ingestion, dermal contact with the body, and inhalation of gas and particle phase.
- Many of the results can not only be used for policy making in the Netherlands but also be informative for other countries with similar agricultural practices and topographies. Moreover, this protocol—with the description of the study design—can serve as a basis for studies in countries with different agricultural practices but common goals.

Methods

Study Design

The OBO study started in January 2016. Enrollment and sample collection were performed in 2016 and 2017. Sample and data analysis were done almost in parallel from mid-2017 to mid-2019. The study focused on flower bulb cultivation and downward spray applications.

An exposure assessment strategy was developed to include personal sampling, environmental sampling, and the collection of contextual information. Additional experimental studies were conducted to generate complementary information on methods of urine collection from non–toilet trained infants [27], the toxicokinetics of human metabolites of pesticides [28], as well as experimental applications to better understand pesticide spray drift and volatilization. The study design is shown in Figure 1. At the start of the OBO study (Module 1), the focus was on identification and selection of pesticides to be analyzed and fields, homes, and participants to be studied (ie, residents living in selected homes). In Module 2, exposure assessment was conducted in and around the homes after one of the selected pesticides was sprayed on a selected field. Methods for diaper sampling and assessing personal pesticide exposure were developed in Modules 3 and 4, respectively. On some of the fields, spray drift experiments (Module 5) and volatilization experiments (Module 6) were conducted. Finally, results from Modules 2, 4, 5, and 6 provided input for Module 7, the modeling of exposure for each of the homes (from Module 1). Each module is discussed in more detail below.
Module 1: Identification and Selection of Pesticides, Fields, Homes, and Participants

Pesticides

In the selection of relevant pesticides to target in our chemical analyses, the main aspects taken into account were (1) information about registration and usage of pesticides on flower bulbs for the year 2015, collected from available data [26] and interviews with growers; (2) existing monitoring data for soil/crops from flower bulb fields; (3) amenability to multiresidue analysis methods; (4) estimated deposition and source strength of emissions from plants and from the top soil layer; (5) estimated dermal exposure and skin absorption potential; and (6) possible exposure originating from other, nonagricultural pesticide use (e.g., food consumption) [29]. Detailed inclusion and exclusion criteria are provided in Figure 2.

Figure 2. Inclusion and exclusion criteria used in the selection of pesticides to be analyzed in environmental and urine samples. LC-MS/MS: liquid chromatography with tandem mass spectrometry.
it covered the largest number of the targeted pesticides. In addition, this method was also more suited to include several relevant degradation products/metabolites. The final selection included 46 prioritized pesticides/metabolites to be measured in air, dust, and soil (see Multimedia Appendix 2).

For the analysis of the urine samples (internal exposure assessment at personal level), the target analyte (biomarker of exposure) in almost all cases was not the parent pesticide but a metabolite formed upon uptake. Of the 46 pesticides selected for environmental measurements, data on their human biomarkers, analytical standards of the potential biomarkers, and methods for their analysis were not yet available. Consequently, as part of the OBO study, data on biomarkers and excretion profiles had to be generated (see Module 4), analytical standards synthesized, and methods for analysis developed. This was a substantial effort and obviously could not be done for all 46 pesticides. For this reason, the assessment of internal exposure was restricted to a subset of 5 pesticides, which should be sufficiently representative to facilitate modeling and extrapolation to other pesticides. Ideally, the pesticides selected for biomonitoring represented the 3 main product types (herbicide, insecticide, and fungicide), different physicochemical properties of the pesticide, and actual spray applications on flower bulbs. We made a short list of 8 prioritized pesticides that were frequently sprayed in bulb fields and could serve as representatives for the whole set and offered good prospects regarding the biomarker analytical challenges. In short, these criteria pertained to factors such as representing different physicochemical properties, pesticide market shares, frequency of application, dosage, vapor pressure, half-life in the environment, and dermal absorption rate. To minimize the influence of dietary contributions on the biomarker levels, we also considered the likelihood of being present in food items.

The 8 selected pesticides were chlorpropham, asulam, fonicamid, acetamiprid, thiacloprid, prochloraz, tebuconazole, and trifloxystrobin. All of these substances were expected to be routinely used by the growers, and most of them, with the exception for chlorpropham, have a low likelihood of dietary exposure compared with other pesticides (see Multimedia Appendix 3). However, as indicated above, because of feasibility constraints, only a maximum of 5 pesticide biomarkers could be analyzed in urine (ie, B in Figure 2 must be equal to 5). Finally, the pesticides (biomarkers) that were selected for biomonitoring were asulam (asulam), carbendazim (methyl 5-hydroxy-2-benzimidazole carbamate [5-HBC]), chlorpropham (4-hydroxychlorpropham-O-sulfonic acid [4-HSA]), prochloraz (2,4,6-trichlorophenol [2,4,6-TCP], and tebuconazole (tebuconazole-1-hydroxy [TEB-OH]).

Fields
Selected fields needed to meet the following criteria: (1) residents’ homes were located in the vicinity (within 250 m) of flower bulb fields; (2) growers had a previously defined cultivation plan; and (3) growers were willing to participate and share their spray plan (including product formulation, amount applied, type of nozzle used, and spraying date and hour) with the research team.

Here, we defined location as a place consisting of one or more agricultural fields, with at least one bulb cultivation and surrounded by homes at different distances from those fields. An evaluation comprising a visit to the locations and a meeting with the growers resulted in the final selection of study locations.

It is important to note that there were other fields, besides the selected fields, within 250 m of participating homes. To account for this, growers of all fields near a home (<250 m) that could potentially influence indoor and outdoor environmental pesticide concentrations were asked to share their spraying schemes. In the case of no collaboration (40%), spraying schemes were generated based on type of bulb, weather conditions, and standard spray schemes of the crop type reported by local expert agronomists.

Homes
Spray applications on a field may expose residents to pesticides through spray drift and volatilization. Homes located within a 50 m distance at the downwind side of the treated field have been described as directly exposed to spray drift [30]. The pesticide deposited on crop and/or soil may volatilize, and this process might affect homes in each direction, especially if they are located within a short distance (ie, up to 250 m) [31]. Therefore, residents living in homes located within 250 m from a selected field were invited to join the study, with, ideally, recruited homes situated at different distances around that field (Figure 3). Control homes were also included in the study. These homes were located in semirurban areas (ie, <1500 residential addresses/km²) that were situated within 20 km from a selected field but did not have agricultural fields within a 500 m distance.
Participants

Before residents were contacted, the study protocol was approved by the Medical Ethical Committee of the University Medical Center Utrecht (Protocol number NL54727.041.15).

Residents were invited via a letter accompanied by a brochure explaining the study. Interested invitees were interviewed by phone to check if they met the following inclusion criteria: (1) having his/her primary place of residence at the preselected location; (2) having sufficient knowledge of the Dutch language and no cognitive impairment and therefore able to complete the administered questionnaires and communicate with the study assistant; and (3) without a diagnosis of kidney or liver disease, as these could change metabolite formation.

Once enrolled in the study, participants were asked about availability and willingness of further household members (partners and children) to participate.

Module 2: Exposure Assessment

For a comprehensive exposure assessment, environmental and urine samples were collected as well as information regarding daily activities, food consumption, building characteristics, and other relevant factors using field forms, questionnaires, and diaries. These data were gathered on two occasions, hereafter referred to as "measuring campaigns": during the pesticide use period (UP) and during the nonuse period (NP).

During the UP, environmental samples from homes were collected after a reported spray event on a selected field (Module 1). Outdoor air was sampled for 7 consecutive days because this is the period of time that we expected to see an influence on concentrations due to spray drift (day 1) and evaporation (days 1 to 7) of pesticides. This expectation was based on detailed model calculations of spraying events (using the models described below in Module 7). During the NP, we only expected background concentrations, and therefore we sampled for a shorter period of time (2 days). Biomonitoring was performed on the same days as environmental samples were collected.

For almost all other environmental samples, namely vacuumed floor dust (VFD), dust from a newly placed clean doormat (DDM), windowsill dust, soil from the garden (if one existed), and soil from the selected field, collection took place at the end of the 7-day and 2-day period, respectively, for the UP and the NP. Additionally, in both the UP and the NP, an electrostatic dust collector (EDC) was placed at the start of the measuring campaign and collected at the end.

Regarding personal sampling, morning urine samples were collected daily for 7 consecutive days and hand wipes were taken on the first day of urine collection.

A measurement campaign was set in motion through a system allowing remote initiation of the air pumps once the grower informed the research team that spraying of at least one of the 8 short-listed pesticides was scheduled to begin. This ensured that our sampling periods were aligned with an actual application.
Module 3: Development of Diaper Sampling
Self-collection of urine by adults and toilet trained children was done using a 1 L measuring cup and 250 mL plastic jars. To determine the best method for urine collection in non-toilet trained infants (aged 0-3 years), four commonly applied methods were evaluated in a pilot study (in a nonclinical setting). The four methods were (1) free catch, (2) a urine collection pad (Hessels & Grob BV), (3) a urine bag (Urinocol Pediatric, Braun), and (4) a disposable polyacrylate diaper (Pampers Baby Dry size 3, Procter & Gamble). The study examined the success scores of sample collection by parents/caretakers and acceptance scores by infants and parents/caretakers. The most successful and best-accepted method—and also the one that collected a sufficient urine volume (>5 mL) to allow biomarker analyses—was the disposable diaper [27]. This was the method used for urine collection in non-toilet trained infants in this study.

Module 4: Identification and Toxicokinetics of Urinary Biomarkers
For most pesticides, metabolism in humans is unknown, and the only available data are derived from animal studies. For urine biomarker analysis, knowledge about the most suitable (specific and sensitive) human biomarker was needed in addition, to link urinary concentrations (internal exposure) to external exposure, knowledge of toxicokinetics and urinary excretion profiles was needed. For this, human volunteer studies were set up for each of the 5 pesticides selected for biomonitoring. Each study involved two independent administrations of the pesticide—one oral and one dermal (2 weeks apart)—to a group comprising 3 males and 3 females. Individual urine samples were collected for 48 hours. First, a biomarker screening was performed for composite urine samples using liquid chromatography–full-scan high-resolution mass spectrometry. For the most suitable biomarker tentatively identified, the analytical standard and its isotopic analog were purchased. In most cases, this required custom synthesis, especially for the isotopic analogs. Following full conformation, dedicated methods for analysis of each biomarker were developed and validated, and all individual samples from each of the volunteers were analyzed. In this way, data on toxicokinetics were generated and conversion factors were derived [28]. The conversion factors were used to estimate pesticide uptake (Module 7) based on measured urinary biomarker concentrations (Module 2).

Module 5: Spray Drift Experiments
Spray drift models were developed previously to estimate the environmental fate (ie, spray drift deposition at ground surface and airborne) of pesticides near application areas [32]. However, since residential exposure was not considered during the development of these models, there were knowledge gaps in predicting residential exposure, especially at larger distances (>15 m) from the field and at greater heights (>3 m). To address these gaps, experimental studies were carried out on 6 agricultural fields to study spray drift at longer distances (5 m to 50 m) and greater heights (up to 10 m) as well as the effect of physical barriers. The application techniques for downward spraying were similar to those used in practice. The types of nozzles used were a TeeJet XR11004 (TeeJet Technologies) and agrotop TDXL11004 (agrotop GmbH). These are respectively standard and 90% drift-reducing flat-fan nozzles [33].

For ethical and practical reasons, measurements were performed using a fluorescent tracer instead of pesticides. Experiments were repeated using the aforementioned nozzle types as well as with varying foliage coverage on the field (ie, bare ground to full crop). The results from these studies helped to calibrate the spray drift model, which provided output for use in modeling exposures (Module 7).

Module 6: Volatilization Measurements
Pesticide volatilization experiments were also conducted. Two experimental sites were selected based on the defined field and crop types (eg, type of flower bulb) selected in Module 1. In the selected locations, rates of pesticide volatilization from the treated crops and influencing factors were measured on the day of pesticide application and several times during the first week after application. This was achieved by combining measurements of pesticide concentration gradients and on-site meteorological observations, including measurements of turbulence intensity. In addition, the pesticide residues on leaves were determined. Results of these measurements were used to test the volatilization model [34], which provides hourly emissions from fields due to volatilization for use in modeling exposures (Module 7).

Module 7: Modeling of Exposures
In order to select models suitable for assessing the exposure of residents living near fields where pesticides are intensively used, a screening of different models was conducted [35]. The most suitable models were combined into a deterministic modeling framework (Figure 4). Selected models were calibrated with results from measurements and experimental studies (Modules 5 and 6; Figure 4A). Model estimates were verified by comparing predicted concentrations in different media (eg, air, dust, soil) with concentrations measured in and outside homes (Module 2; Figure 4B). Next, the deterministic models were used to estimate pesticide exposure of residents living within 250 m of fields where spraying applications occur (Module 7). In this module, the contributions of different exposure routes to total internal exposure were investigated. In addition, different factors (eg, personal pesticide use, time spent indoors) that might influence personal exposure were incorporated via statistical modeling techniques.
Figure 4. Deterministic modeling framework. (A) Models were calibrated with results from measurements and experimental studies. (B) Modeling steps were verified by comparing predicted and measured concentrations.

Sample Size
To determine our sample size, we performed a power calculation based on National Health and Nutrition Examination Survey data (US Centers for Disease Control and Prevention) on urinary 3-phenoxybenzoic acid (metabolite of pyrethroid pesticides) concentrations. With 200 residents, we estimated that we would reach 80% power at an $\alpha=0.05$ level to detect a 40% to 100% difference between background levels (mean 0.292 $\mu$g/L, SD 0.26 $\mu$g/L) and exposed individuals, assuming an exposure prevalence of 50% and 100%, respectively. Therefore, we aimed to include 200 residents (roughly 100 homes) in our study.

Data Collection

Measurements
As mentioned in Module 2, different types of environmental and personal samples were collected during this study. All of the different sample and collection procedures are summarized in Table 1. Environmental samples were transported to the laboratory within 48 h after sampling. Air sampling cartridges and dust were stored at 4°C until analysis, while soil and crop samples were stored at –18°C. Analysis of the environmental samples was based on existing methods already available in the consortium laboratories. The methods were slightly adapted to include all 46 pesticides selected and revalidated according to SANTE/11813/2017 (currently SANTE/12682/2019). The latter included establishment of recovery, repeatability, selectivity, and limit of quantification (LOQ; defined as the lowest successfully validated level). For air analysis, the glass fiber filter (trapping airborne particles) and the Amberlite XAD-2 adsorbent (Sigma-Aldrich, Inc; trapping gas-phase pesticides) were combined and extracted by accelerated solvent extraction using acetonitrile/methanol. After evaporative preconcentration, the pesticides were analyzed by LC-MS/MS. LOQs were 0.01 ng/m³ for most pesticides. For the analysis of household dust, soil, and crops, extraction was done by a mixture of water/acetonitrile, followed by a salt-induced phase partitioning technique (QuEChERS) [36,37]. The organic phase was analyzed by LC-MS/MS. LOQs were 1 $\mu$g/kg for most pesticides.

Urine samples were transported to the laboratory within 24 h after collection, aliquoted, and then stored at –18°C. For each biomarker, a different dedicated method was developed to obtain optimum performance, which was validated and then applied to sample analysis. In all cases, the isotope analog of the biomarker was added as an internal standard to the urine aliquot to be analyzed (1 mL to 5 mL) at the start of the sample preparation. For biomarkers of tebuconazole (TEB-OH), prochloraz (2,4,6-TCP), and thiophanate-methyl/carbendazim (5-HBC), an enzymatic deconjugation step was performed. Biomarkers of chlorpropham (4-HSA) and asulam (parent compound) were analyzed directly. For the other biomarkers, extraction/cleanup involved either solid phase extraction or a liquid-liquid partitioning step (QuEChERS-based), followed by an evaporative concentration step. Analysis of the extracts was done by LC-MS/MS under optimized conditions for the respective biomarkers. LOQs were as follows: 0.1 ng/mL for asulam, 4-HSA, and TEB-OH; 0.05 ng/mL for 5-HBC; and 0.25 ng/mL for 2,4,6-TCP.

Not all of the collected samples were analyzed; the unanalyzed samples were kept under appropriate storage conditions for future analysis. Samples to be analyzed were selected based on the location of the home—to guarantee a good distribution of distances of homes to the selected field—and on the wind direction during the application. This resulted in groups of homes per different distances (ie, homes between 0-50 m, 50-150 m, and 150-250 m) both downwind and upwind of the fields where applications took place. All collected samples from the selected homes were analyzed. For the remaining homes, only DDM was analyzed, providing us with an idea of the distribution of indoor dust concentrations at all locations.
Table 1. Samples and collection methods used in the OBO study ("Research on exposure of residents to pesticides").

<table>
<thead>
<tr>
<th>Sample</th>
<th>Collection method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outdoor and indoor air</td>
<td>Air was sampled through a standard PM10 inlet and drawn through a glass fiber filter and a tube containing XAD-2 adsorbent (Amberlite XAD-2; Sigma-Aldrich, Inc.).</td>
</tr>
<tr>
<td>Vacuumed floor dust</td>
<td>A dust sampling sock (Allied Filter Fabrics Pty Ltd) was attached to the hose of a vacuum cleaner to sample for 5 minutes on 4 m² of carpet or 6-8 m² of smooth floor.</td>
</tr>
<tr>
<td>Dust from doormat</td>
<td>A clean doormat was deployed for 1 week and then vacuumed on arrival to lab facilities.</td>
</tr>
<tr>
<td>Soil from each field and each residential garden</td>
<td>Five uncovered areas of soil were randomly selected and approximately 150 g to 250 g of topsoil were collected per area and combined into a single aggregate soil sample.</td>
</tr>
<tr>
<td>Tank mix sample</td>
<td>Duplicate tank mix samples of the spraying liquid were taken directly before and immediately after the spray event in all selected fields (Module 1). Aliquots of the tank samples were stabilized with methanol.</td>
</tr>
<tr>
<td>Windowsill dust</td>
<td>Clean wipes were used to collect dust accumulated on windowsill surfaces.</td>
</tr>
<tr>
<td>Electrostatic dust collector (EDC)</td>
<td>EDCs were deployed inside each home at the start of the study and collected at the end of the sampling period.</td>
</tr>
<tr>
<td>Urine b</td>
<td>Spot samples were collected from all participants, except for non–toilet trained toddlers.</td>
</tr>
<tr>
<td>Hand wipe b</td>
<td>The hand wipe consisted of a facial tissue premoistened with 3 mL of a 50% water/50% ethanol solution.</td>
</tr>
</tbody>
</table>

Analyses were performed using targeted liquid chromatography with tandem mass spectrometry for the main biomarkers of 5 preselected pesticides. Personal sampling was performed.

Questionnaires and Diaries

For each home, the research assistant filled in a field form with building characteristics (see Multimedia Appendix 4), with data on the type of flooring, age of the building and materials used to build it, volume and area, number of floors, type of ventilation system, type of heating, and possible air leakages (eg, cracks). Each participant also completed a questionnaire and diary (see Multimedia Appendix 4) on personal characteristics, socioeconomic position, presence and type of pets, use of medications, education level, type of work/education, whether shoes were worn indoors, and personal use of pesticides. Parents were asked to fill in the questionnaires for their children. Questionnaires were completed before the measurement campaigns started. During measurement campaigns (ie, during both the UP and the NP), participants filled out a daily diary on food intake, hours spent at home and/or elsewhere, and personal use of chemicals, biocides, or pesticides. Via an additional short questionnaire, we checked if items on the original personal questionnaire had changed during the measurement campaigns.

Data Management

All data collected from the field study were transferred to the OBO data manager at Utrecht University. Entry of the collected questionnaire data was done using the Castor EDC interface, making our data storage compliant with relevant regulations, such as the General Data Protection Regulation 2016/679, International Organization for Standardization (ISO) 27001, and ISO 9001 [38]. For diaries, we used a tailor-made data entry program. The entry was done in duplicate and then checked against each other (100% check). A third person looked at the differences, and if errors were found, records were rechecked against the original hard copies. Once completed, pseudonomized data were used for analyses.

Ethics: Stakeholder Engagement and Dissemination

The OBO study was commissioned by the Netherlands National Institute for Public Health and the Environment. It was conducted by a consortium of Dutch institutes including Utrecht University, the Netherlands Organization for Applied Research, Wageningen University & Research, Radboud University Medical Center, consulting and communication agency Schutteelaar & Partners, CLM Research and Advice, and Professor PJJ Sauer. The research proposal has been reviewed by a panel of 16 international experts. During the preparation and execution of the OBO study, a stakeholder group advised on all research, communication, and ethical aspects. This group consisted of representatives of public sector policy makers, researchers, the private sector, nongovernmental organizations, and citizens.

Results

Identification and Selection

Of the contacted growers of possible selected fields, 17% (12/70) participated in the study. Nine fields were included, encompassing spraying at different crop stages, variability for different meteorological parameters (such as temperature and wind speed), and application of 20 different pesticides (the majority were fungicides [9/20, 45%]). Some fields were sprayed more than once during the UP, which enabled us to follow a total of 14 different primary spraying applications at our selected fields. A total of 80 homes around the selected fields and 16 control homes were included in the study, with a total of 192 participants, of which 39 were younger than 18 years old. An overview is provided in Multimedia Appendix 5.
participate. In total, the residents of 80 homes around the selected fields responded and were included in the study (response rate of 4.5%), while the residents of 16 control homes responded and were included (response rate of 3.3%). We were able to have a good spatial distribution of homes around the selected fields: 25% (20/80) of homes were located within 50 m, 43% (34/80) were between 50 m and 150 m, and 33% (26/80) were between 150 m and 250 m from the selected fields. Of the 192 participants, 164 were residents living within 250 m of a selected field. In this group, slightly more than one-half of the participants were female (89/164, 54%) and the average age at participation was 44 years (range 2 to 88 years). Of the 28 participants living in control homes, slightly more than one-half were male (16/28, 57%) and the average age at participation was 50 years (range 12 to 76 years).

We selected 46 active ingredients of pesticides (see Multimedia Appendix 2) for environmental analysis. For biomonitoring, we selected the following 5 pesticide biomarkers: asulam (asulam), carbendazim (5-HBC), chlorpropham (4-HSA), prochloraz (2,4,6-TCP), and tebuconazole (TEB-OH). Carbendazim was not on the initial short list; analyses of indoor dust in the initial phase of the project led to the choice to include this substance in our selection because it was detected in almost all indoor dust samples, often in co-occurrence with thiophanate-methyl. Both are fungicides. Carbendazim, which is no longer registered, arises from the use of thiophanate-methyl, which transforms into carbendazim both in the environment and upon uptake by humans. Thiophanate-methyl has no field spray application in bulb fields but is used for bulb disinfection. It might be emitted from the bulb disinfection site and/or end up in the field upon planting of the bulbs.

**Exposure Assessment**

In total, we collected 969 outdoor air samples, 49 indoor air samples, 224 VFD samples, 221 DDM samples, 238 soil samples from residents’ gardens, 27 soil samples from the application fields, 2054 morning urine samples, 431 daytime urine samples, 112 hand wipes, and 91 tank mix samples. We analyzed approximately one-half of all collected samples. These consisted of 628 outdoor air samples, 43 indoor air samples, 128 VFD samples, 170 DDM samples, 124 soil samples from residents’ gardens, 20 soil samples from the application fields, 791 morning urine samples, and 311 daytime urine samples. All collected hand wipes and tank mix samples were analyzed. The spray events and the respective applied pesticides and tank mixtures are shown in Table 2. The modeling framework was used and verified in all presented locations under several different meteorological conditions (see Multimedia Appendix 5).
<table>
<thead>
<tr>
<th>Location, selected field size, measurement campaign, and pesticide sprayed</th>
<th>Grower’s self-reported dosage (kg/ha)</th>
<th>Measured dosage (kg/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A: 2.45 ha</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UP1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folpet</td>
<td>0.23</td>
<td>nd&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mancozeb</td>
<td>1.50</td>
<td>nd</td>
</tr>
<tr>
<td>Tebuconazole&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Thiacloprid&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>B: 2.29 ha</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UP1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flonicamid&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>Fluopyram&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Trifloxystrobin&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>C: 2.00 ha</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UP1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorpropham&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.80</td>
<td>0.76</td>
</tr>
<tr>
<td>Pendimethalin&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.80</td>
<td>nd</td>
</tr>
<tr>
<td><strong>UP2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mancozeb</td>
<td>1.88</td>
<td>nd</td>
</tr>
<tr>
<td>Tebuconazole&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>D: 1.09 ha</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UP1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorpropham&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.80</td>
<td>0.88</td>
</tr>
<tr>
<td>Pendimethalin&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.80</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>UP2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorothalonil</td>
<td>0.50</td>
<td>nd</td>
</tr>
<tr>
<td>Esfenvarelate</td>
<td>0.01</td>
<td>nd</td>
</tr>
<tr>
<td>Mancozeb&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1.24</td>
<td>nd</td>
</tr>
<tr>
<td>Prochloraz&lt;sup&gt;c,e&lt;/sup&gt;</td>
<td>0.16</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>E: 4.58 ha</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UP1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetamiprid&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>Esfenvarelate</td>
<td>0.01</td>
<td>nd</td>
</tr>
<tr>
<td>Mancozeb</td>
<td>1.50</td>
<td>nd</td>
</tr>
<tr>
<td>Mepanipyrim&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>UP2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambda-cyhalothrin&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.01</td>
<td>nd</td>
</tr>
<tr>
<td>Mancozeb</td>
<td>1.50</td>
<td>nd</td>
</tr>
<tr>
<td>Flonicamid&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Tebuconazole&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.08</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Discussion

There is ongoing concern in the Netherlands regarding the use of pesticides and their potential impact on the environment and human health. In the last decade, several initiatives and regulations have been implemented to reduce the use of pesticides and to reduce the emission of pesticides during (spray drift) and after (volatilization) applications. However, there remains a lack of information on exposure of residents to pesticides coming from agricultural fields. The OBO study was designed to provide comprehensive insight into the exposure of residents and contributing exposure routes.

Strengths

The design of the OBO study has many strengths. We collected multiple sample types from various matrices in both UPs and NPs. This allowed us (1) to compare exposed locations with control locations in both UPs and NPs; (2) to compare
environmental concentrations among exposed locations by UP and distance to fields; (3) to study the interrelationships between concentrations in various matrices (eg, air and dust); (4) to compare biomonitoring results between exposed and control subjects; (5) to relate biomarker levels to environmental concentrations; and (6) to use measurements for model calibration and verification.

Using a single LC-MS/MS method, we managed to determine 46 active substances in the environmental samples. This group of substances covered approximately 60% of the different pesticides registered in the spraying records around the selected fields. The inclusion of 46 different substances allowed us to analyze substances applied in the selected fields, substances applied in other fields in the vicinity, and pesticides with no recorded use in the area, which enabled us to compare patterns between these different use categories.

An emphasis of our study was on modeling of the exposure of residents to pesticides. This resulted in a framework of models that may be useful to also estimate exposure from substances and mixtures that were not included in our study.

Limitations

One limitation of the study is that participating growers were not blinded (ie, they were informed of the research aim); therefore, one could hypothesize that growers sprayed only under certain conditions (eg, if the wind was blowing away from residential homes or at a very low speed). Of course, in the end, growers will spray when they need to so as to avoid cultivation loss. To account for this, we collected information from multiple applications that occurred not just on the selected field but also on surrounding fields.

Another limitation is that our study focused only on exposure as a result from downward spraying. Information regarding the degree of exposure to pesticides of residents living near crops where sideways or upward spraying techniques are used (such as in fields of fruit trees) is still lacking in the scientific literature and needs to be assessed in future studies. Here, a study design similar to the one used in the OBO study can be used.

Finally, it is important to add that given the low participation rates, the homes and residents included in our study might not be representative of the population living in the vicinity of agricultural fields.

Lessons From the OBO Study

There are several lessons learned from the OBO study, some related to co-creation (see Textbox 2) and others related to practical aspects (see Textbox 3). We feel that these are relevant to the research community and might help future projects in tackling these a priori.

Textbox 2. Lessons related to co-creation.

- It was extremely important to have a clear line of communication with the stakeholders and involve them in both the design and the study period. This allowed us to address their concerns upfront. Communication was maintained via presential meetings, and the outcome of the meetings was then transmitted through the entire OBO consortium. We noticed that it was very important to make the aim of the project clear from the start and to check if all stakeholders understood the main goals (ie, managing expectations).
- The collaboration with the branch organizations was also very important as they assisted with the recruitment of growers.
- Information events proved very helpful for dissemination and discussion of results in both the local communities and with the growers.
### Textbox 3. Lessons related to practical aspects.

#### Participation and selection

- Obtaining the participation of selected households (at different distances and in different directions) around selected fields with a diverse population (i.e., in age and sex) proved to be a difficult task. Although we achieved a good spatial distribution of homes per distance to selected fields overall, in some locations more than 50% of the homes were located in one of the distance categories (i.e., <50 m, 50-150 m, or 150-250 m). Ideally, for each location we would have had one-third of the homes in each of the distance categories. As well, additional requirements, such as multiple persons within a household and the presence of children, made selection of appropriate households challenging.

- The recruitment of growers proved difficult both because of the underlying needs of the project regarding data (e.g., a spraying schedule) and because of pressure within the agricultural sector (e.g., “is the outcome of the project going to affect me and other growers?”). However, once growers were enrolled in the study, they continued until the end. As mentioned above, the involvement of key branch organizations was very important in this step.

- In the selection of urinary biomarkers, we selected pesticides with a lower exposure background via food intake to make the environmental exposure signal more clear. This was done based on information available in the literature, but it was difficult to achieve high specificity.

#### Periodicity of spraying applications

- Regarding the active air sampling, the logistics were complex, as the exact timing of application was often unknown. Our solution was to install all equipment and wait until the application; however, this created some downtime periods where there were not enough measuring instruments available. Aside from changes in the intended date of application because of changes in meteorological conditions, regular communication with the grower regarding the date of intended application was very useful for the field work planning.

#### Analytical standards

- Synthesis of the analytical standards (and isotope labels) took considerable amount of time and caused a delay in analysis of the urine samples. However, it was difficult to do this in another way. We had to know what was actually sprayed and what was found in the environment; only then could we finalize the selection of the 5 pesticides, start volunteer studies, and, finally, get the biomarkers synthesized. This is important to take into account when setting up a new study.

- In retrospect, the dust samples provided a lot of valuable information regarding presence of various pesticides in the environment. Thus, it would have been useful if we had done a full-scan prescreen of household dust in houses (and fields) of candidate growers and residents’ homes. For that, no ethics approval was needed, so we could have done that at a very early stage of the study, during the time when we were working on pesticide selection.

#### Exposure of residents to pesticides and communication of results

- It is important to take into account that we might have only captured a few different exposure scenarios by doing field work. We were constrained by existing meteorological conditions and by the applications that occurred within that time window. As a solution, we used the developed modeling framework to simulate realistic worst-case scenarios by looking into long-term meteorological ranges and different applications settings.

- At the beginning of the project, we promised participants that they would receive feedback on their results, but given the abovementioned time delays, this took a longer time, which resulted in a frustrating process for participants. For future projects, we recommend informing participants a priori of possible delays that might occur.

- Given the very high sensitivity of the methods used, detected exposures may still translate into very low absolute exposures. Therefore, results need to be carefully communicated in order to prevent possible misunderstandings.

### Conclusion

The OBO study can shed light on current and future questions through the materials collected, methods developed, and wealth of data generated. These can, for example, be used for testing model improvements, to put results of other exposure experiments into context, or to develop new hypotheses, thereby also setting the stage for future collaborations.

### Acknowledgments

This study was funded by the Dutch Ministry of Infrastructure and Water Management and the Ministry of Economic Affairs and Climate Policy. We would like to thank all participants and growers for their willingness to participate in this study, the stakeholders for providing continuous feedback, and all the parties involved in the OBO study group.

We cordially invite other researchers to propose noncommercial research based on the available data in OBO or requests for additional chemical analyses with associated funding. Any such requests can be submitted to exposome.office@uu.nl with subject: OBO-research.

### Conflicts of Interest

None declared.
Multimedia Appendix 1
List of most relevant pesticides in flower bulb cultivation in the Netherlands.
[DOCX File, 22 KB - resprot_v10i4e27883_app1.docx ]

Multimedia Appendix 2
List of 46 pesticides targeted for analysis of environmental samples and their main chemical properties.
[DOCX File, 23 KB - resprot_v10i4e27883_app2.docx ]

Multimedia Appendix 3
Short list of 8 pesticides selected for biomonitoring.
[DOCX File, 22 KB - resprot_v10i4e27883_app3.docx ]

Multimedia Appendix 4
OBO questionnaire, diary, and field form (translated).
[DOCX File, 36 KB - resprot_v10i4e27883_app4.docx ]

Multimedia Appendix 5
Selected fields, measurement campaigns, meteorological conditions, and participating homes and residents.
[DOCX File, 26 KB - resprot_v10i4e27883_app5.docx ]

Multimedia Appendix 6
External peer reviewers report from the National Institute for Public Health and the Environment (RIVM). This report comprises the comments of 16 external reviewers. The comments are in Dutch.
[PDF File (Adobe PDF File), 422 KB - resprot_v10i4e27883_app6.pdf ]

References


22. de Vreede JA, Brouwer DH, Stevenson H, van Hemmen JJ. Exposure and risk estimation for pesticides in high-volume directed spray applications in field crops. URL: https://edepot.wur.nl/448381 [accessed 2020-11-14]


25. van de Zande JC, ter Horst MMS. Crop related aspects of crop canopy spray interception and spray drift from downwind directed spray applications in field crops. URL: https://edepot.wur.nl/514310 [accessed 2021-04-14]


Abbreviations
2,4,6-TCP: 2,4,6-trichlorophenol
4-HSA: 4-hydroxychlorpromam-O-sulfonic acid
5-HBC: methyl 5-hydroxy-2-benzimidazole carbamate
DDM: dust from a newly placed clean doormat
EDC: electrostatic dust collector
ISO: International Organization for Standardization
LC-MS/MS: liquid chromatography with tandem mass spectrometry
LOQ: limit of quantification
NP: pesticide nonuse period
OBO: Research on exposure of residents to pesticides in the Netherlands
TEB-OH: tebuconazole-1-hydroxy
UP: pesticide use period
VFD: vacuumed floor dust

Edited by T Derrick; submitted 12.02.21; this is a non–peer-reviewed article; accepted 19.02.21; published 28.04.21.

Please cite as:
Pesticide Exposure of Residents Living Close to Agricultural Fields in the Netherlands: Protocol for an Observational Study
JMIR Res Protoc 2021;10(4):e27883
URL: https://www.researchprotocols.org/2021/4/e27883
doi:10.2196/27883
PMID:33908892

©Daniel M Figueiredo, Esmeralda J M Krop, Jan Duyzer, Rianda M Gerritsen-Ebben, Yvonne M Gooijer, Henk J Holterman, Anke Huss, Cor M J Jacobs, Carla M Kivits, Roel Kruijne, Hans J G J Mol, Arné Oerlemans, Pieter J J Sauer, Paul T J Scheepers, Jan C van de Zande, Erik van den Berg, Marcel Wenneker, Roel C H Vermeulen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

Structure and Functioning of Acute Inpatient Psychiatric Units in Spain: Qualitative Study

Roberto Rodriguez-Jimenez, MD, PhD; Iluminada Corripio, PhD; Ricardo Campos, PhD; Mario Páramo, MD; Manuel Franco-Martín, MD, PhD; Estefanía Segura, MD; Sergio González, MD; José Martínez-Raga, MD, PhD

1Department of Psychiatry, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain
2Biomedical Research Networking Centre in Mental Health (CIBERSAM), Madrid, Spain
3Faculty of Medicine, Universidad Complutense de Madrid, Madrid, Spain
4Department of Psychiatry, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
5Department of Psychiatry, Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain
6Department of Psychiatry, Hospital de Conxo, Universidad de Santiago, Santiago de Compostela, Spain
7Department of Psychiatry, Jefe del servicio de psiquiatría del Hospital Universitario Rio Hortega de Valladolid, Valladolid, Spain
8Department of Psychiatry, Complejo Asistencial de Zamora, Zamora, Spain
9Department of Psychiatry, Hospital General Universitario de Ciudad Real, Ciudad Real, Spain
10Department of Psychiatry, Hospital Universitario Virgen Macarena, Sevilla, Spain
11Department of Psychiatry, Faculty of Medicine, University of Seville, Sevilla, Spain
12Department of Psychiatry, Hospital Universitario Doctor Peset, Valencia, Spain

Corresponding Author:
Roberto Rodriguez-Jimenez, MD, PhD
Department of Psychiatry
Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12)
Av de Córdoba km 5,400
Madrid, 28041
Spain
Phone: 34 91 390 85 36
Email: roberto.rodriguez.jimenez@gmail.com

Abstract

Background: As a consequence of the decentralization of health care provision to the different Regions (called Autonomous Communities) in Spain, different health care models and resources have been developed for psychiatric patients. It would be very useful to obtain comprehensive and comparative data on health care models, resources, and activity of acute inpatient psychiatric units (AIPUs) as a key part of mental health systems.

Objective: The aim of this study was to determine the current state of AIPUs in Spain through a national scorecard that allows the current situation to be visualized in terms of resources, processes, and outputs.

Methods: A 104-item online questionnaire was sent to all the AIPUs of the different Regions in Spain. It was divided into 11 sections, including data on the resources, processes, and outputs of the AIPUs plus general data, an indicator dashboard, and good practices.

Results: The questionnaire was completed by 60.0% (117/195) of the AIPUs invited to participate. The information collected has allowed us to obtain a detailed snapshot of the current situation of AIPUs in Spain at the levels of infrastructure and material resources, staffing, organization and activity of the units, coordination with other units, guidelines, processes and protocols used, participation and communication with patients and their families, teaching activity, and research linked to the units.

Conclusions: This project aimed to help understand the general situation of AIPUs in Spain and its different Regions, contribute to enhancing the benchmarking and harmonization among Spanish Regions, and provide data for future comparisons with other countries.

International Registered Report Identifier (IRRID): RR1-10.2196/26214

(JMIR Res Protoc 2021;10(4):e26214) doi:10.2196/26214

https://www.researchprotocols.org/2021/4/e26214
KEYWORDS
acute inpatient psychiatric units; organization; resources; scorecard; Spain

**Introduction**

In Spain, as a consequence of the decentralization set out in the Spanish Constitution of 1978, the jurisdiction for health care provision was transferred to each of the 17 Regions in Spain (known as Autonomous Communities) to constitute regional health care services. These act as the administrative managerial structure that integrates all the health centers, services, and establishments of each Region, as well as the provincial and city councils and any other interregional territorial administration.

Thus, since 1981, health care functions and services have been transferred from the central government to the different Spanish Regions. The first Region to have these functions transferred was Catalonia in 1981, followed by Andalucía in 1984, and eventually the Region of Madrid, where the transfer process was finalized in 2002. As a result, despite sharing a model and common guideline, there are differences both in terms of structure as well as the way in which the different health care systems function within each Region.

This decentralization of health care services over 2 decades has resulted in the coexistence of different models within the Public Health Care System, despite the fact that the Spanish General Health Law defines common guidelines. Thus, in practice, the transfer process has introduced its own characteristics and differential traits in the implementation of this law in each Region. In this regard, in 1987, the Interterritorial Board of the Spanish National Health Care System was created, to represent and coordinate the services provided by all the Regional Health Care Services. This Board was defined as the “permanent body of coordination, cooperation, communication, and information of Health Services” between each other and with the State Administration, with the aim of “promot[ing] the cohesion of the [Spanish] Health System through an effective guarantee of the rights of citizens throughout the country” [1].

The specialty of psychiatry has not been unaffected by these changes, and currently there are inter-Regional differences in the structure and care of psychiatric patients. This fact is understood by psychiatry professionals in Spain, who have stated in different forums and meetings that there is interest in generating a comprehensive overview of psychiatric care models and the performance of acute inpatient psychiatric units (AIPUs) in Spain, integrating the different operating models in the various Spanish Regions.

The present study was therefore proposed with the aim of understanding the current state of AIPUs in Spain through a scorecard that allows the current situation to be visualized in terms of resources, processes, and outputs. It will permit intra- and inter-Regional analysis of unmet needs, taking into consideration the idiosyncrasies of each Region. Moreover, by comparing the strengths and weaknesses among Regions, benchmarking and therefore harmonization regarding Spanish AIPUs could be proposed to health care planners. Finally, the results of this study could provide data for future comparisons with other countries.

An AIPU is defined as the hospital unit where psychiatric treatments are administered that require short-term inpatient admission (in general, less than 1 month). Admission of a patient to an AIPU is indicated in situations where it is not possible to provide outpatient care for an acute psychiatric process or for the exacerbation of a chronic process [2]. In line with this, the 1983 Mental Health Act for England and Wales defines short hospital stays as those that are less than 28 days [3]. The World Health Organization Mental Health Atlas 2017 [4] defines psychiatric hospital rooms in general hospitals as units that provide hospital health care for the management of patients with acute mental problems and the period of stay is generally short (weeks to months).

**Methods**

**Questionnaire**

A Scientific Committee comprised of professionals who understand the functioning of the AIPU in the different representative Spanish Regions was appointed. A questionnaire consisting of 104 questions was designed to be answered online by the staff member responsible for the AIPU in the different Regions. As a basis for its design, a literature review was carried out to identify the main resources, processes, and results indicators for AIPUs in Spain and at the international level. The search in PubMed was performed using the terms: “acute inpatient psychiatric units” AND (“quality” OR “indicators” OR “mental health services” OR “measuring” OR “improvement”). Indicators and scorecards for results monitoring and assessment included in the mental health plans of the different Spanish Regions, clinical management plans and activity reports of psychiatry services, and recommendations and specialized publications on management systems were also reviewed. Based on the results of this review [5-26], a first draft of the survey was drawn up using the scorecards and indicators of resources, processes, and results in AIPUs. Finally, this first draft was reviewed by the Scientific Committee, who selected what they considered to be the more important variables and indicators. The final questionnaire was included in the 123 Form Builder online survey platform [27]. The online version was also reviewed by the Scientific Committee to ensure that it was easy to understand and complete.

**Data Collection**

The data collection period was open for more than 6 months, between July 2018 and January 2019. At beginning of this period, emails were sent out to all the designated heads of the AIPUs to introduce the study and to provide access to a link to the questionnaire. In order to obtain the most complete responses possible, both in terms of the participating centers as well as the percentage of questions answered per questionnaire, a total of 10 general reminders were sent out via email to the units that had not started or completed the questionnaire. After 4 weeks,
personalized emails were sent out to the person in charge of the unit whenever a response was not obtained.

The online questionnaire sent out to the inpatient units of the different Spanish Regions was divided into 11 sections, including data on the resources, processes, and results of the AIPU, plus a General section asking for general and complementary information about the AIPU, with the aim of collecting, in a structured manner, the most complete set of information on the units (see Table 1).

### Table 1. Survey sections and number of questions.

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Number of questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number of beds</td>
<td>The number of beds available per 100,000 inhabitants</td>
<td>2</td>
</tr>
<tr>
<td>2. Multidisciplinary team of the AIPU²</td>
<td>The number of staff available in each AIPU</td>
<td>9</td>
</tr>
<tr>
<td>3. Infrastructure and AIPU models</td>
<td>The architectural structure, resources and the AIPU model: locked or nonlocked</td>
<td>19</td>
</tr>
<tr>
<td>4. Profiles of patients admitted to AIPUs</td>
<td>Age, patients over 65 years old, patients with cognitive impairment, and patients under 18 years old</td>
<td>8</td>
</tr>
<tr>
<td>5. Access Organization of Care</td>
<td>The route of referral to the AIPU, patient distribution, use of scales, questionnaires, clinical guidelines, and standardized operating procedures</td>
<td>14</td>
</tr>
<tr>
<td>6. Specific interventions offered by the AIPU</td>
<td>Specific interventions such as: medical interdisciplinary consultations, electroconvulsive therapy, or transcranial magnetic stimulation</td>
<td>5</td>
</tr>
<tr>
<td>7. Integration and coordination of the AIPU</td>
<td>The integration and coordination between the professionals that make up the AIPU and also with other hospital services, as well as other levels of care</td>
<td>1</td>
</tr>
<tr>
<td>8. Participation of people with mental health problems and their relatives</td>
<td>The participation of the patients and their relatives in the decision processes and the use of satisfaction questionnaires for patients and relatives</td>
<td>5</td>
</tr>
<tr>
<td>9. Activity of the AIPU</td>
<td>Markers of activity such as total admissions per 100,000 inhabitants, ratio of total planned vs emergency admissions, ratio of voluntary vs nonvoluntary admissions, percentage of admissions per diagnostic group, mean stay, readmissions (at 7 and 30 days), voluntary discharges</td>
<td>6</td>
</tr>
<tr>
<td>10. Training</td>
<td>Number of staff taking part in teaching activities, as well as the number of medical and nursing students</td>
<td>3</td>
</tr>
<tr>
<td>11. Research</td>
<td>The participation of the AIPU in public and private research, development, and innovation</td>
<td>4</td>
</tr>
<tr>
<td>General and complementary information</td>
<td>Hospital type, participant profile, quality indicators, and best practices</td>
<td>28</td>
</tr>
</tbody>
</table>

²AIPU: acute inpatient psychiatric unit.

The survey included a combination of dichotomous (yes/no) and multiple-choice questions, questions asking for specific data or figures, and open questions requiring free-text responses. The results of the survey are presented as mean and standard deviation for quantitative results and as percentages in the case of qualitative variables. Data for Spain as a whole and for each Spanish Region are presented.

### Processes and Outputs

To obtain the profiles of patients admitted to AIPUs [25], data were collected on patient age, percentage of patients over 65 years old, percentage of patients with cognitive impairment, and the percentage of patients under 18 years old.

Regarding access and organization of care [16,17], the following were analyzed: the route of referral to the AIPU (emergency services, mental health center, other services, outpatient clinics, primary care, other), criteria used by AIPU staff to distribute patients who are admitted, use of validated scales and questionnaires in the patient assessment, use of clinical guidelines, and application of standardized operating procedures (such as suicide risk prevention, immobilization, inpatient admission process, electroconvulsive therapy, prevention of unplanned departure from inpatient care, treatment with clozapine).

In addition to the treatment of inpatients, we collected information on the specific interventions offered by the AIPU [13-15]: medical interdisciplinary consultations for hospitalized patients, electroconvulsive therapy, and transcranial magnetic stimulation. Data on access to different pharmacological options for the control of agitation were also collected, as well as access.
to different extended-release antipsychotic drugs for treatment during hospitalization.

The integration and coordination between the professionals that make up the AIPU [11,20], as well as the possible forms of communication, were studied (face-to-face meetings, clinical sessions, telephone, email, electronic health records, videoconference, others). Information was noted on the coordination with other hospital services (eg, internal medicine, neurology, anesthesiology), as well as other levels of care, such as mental health centers, medium-to-long stay units, outpatient hospitals, drug-dependency care networks, and primary care. Finally, data on referrals at discharge were also collected.

The existence of official tools for the participation of the relatives or guardians of patients with mental health issues admitted to the AIPU [6-9], such as personal interviews, informed consent, and written material specifically for good communication with relatives or guardians as well as the patient, was collected. Data on the use of satisfaction questionnaires for patients and relatives were also compiled.

The following activity markers of the AIPU [16-19,21-25] were collected: total admissions to AIPU per 100,000 inhabitants, ratio of total planned vs emergency admissions, ratio of voluntary vs nonvoluntary admissions, percentage of admissions per diagnostic group, mean length of stay, adjusted mean stay index, readmissions to the AIPU (at 7 and 30 days), and voluntary discharges.

Regarding training [17,19,25], data on the number of staff, as well as the number of medical and nursing students, taking part in teaching activities were collected.

To assess research [17,19,25], participation of the AIPU in Research, Development and Innovation (R&D&i) projects and the type of funding were analyzed. We also examined the number of staff from the AIPU who are integrated in established research structures such as the Carlos III Health Institute, Biomedical Research Networking Center for Mental Health Network (CIBERSAM), and Network of Research on Addictive Disorders (Red de Investigación en Trastornos Addictivos). Finally, the different lines of research carried out by the AIPU were collected.

Results

A total of 195 AIPUs were identified in the 17 regions. The designated person in charge of each unit was informed and invited to participate in the study. Finally, information was obtained from 117 AIPUs throughout Spain, providing relevant data on resources and activity; this indicates participation by 60.0% of the identified units. This general response rate, at a 95% confidence level, had a margin of error or confidence interval of 6%. A representation of units from all Regions was achieved (see Figure 1).

Figure 1. Distribution of the contacted acute inpatient psychiatric units (AIPUs) and participation rates per Region.

In most cases, the contact and request for participation were made through the heads of psychiatry services or departments and coordinators of the AIPU; in 70.8% (83/117) of cases, the person who filled out the questionnaire was the coordinator of the AIPU.

According to the functional purpose of the hospital where they are located, 87.2% (102/117) of the participating AIPUs were located in general hospitals, while 12.8% (15/117) belonged to psychiatric hospitals. In terms of the functional dependency of the center, 83.8% (98/117) were in public hospitals. Last, with regards to the hospital size, it should be noted that, although the majority of participating units (54/117, 46%) were located in medium-size hospitals (from 201 to 500 beds, known as group 2 in Spain), the questionnaire was completed by units of hospitals with less than 200 beds (group 1), accounting for 17% (20/117) of all units, and units belonging to hospitals with 500-1000 beds (group 3), accounting for 32% (37/117) of all units. Only 5% (6/117) of the participating units were found in hospitals with more than 1000 beds (group 4; see Figure 2).
Figure 2. Profile of the units that completed the acute inpatient psychiatric unit (AIPU) questionnaire.

**Discussion**

The aim of this study was to understand the current state of AIPUs in Spain through a scorecard that allows the current situation to be visualized in terms of resources, processes, and outputs in the different Spanish Regions. It will permit intra- and inter-Regional analysis and could provide data for future comparisons with other countries.

We estimate that 60.0% (117/195) of all Spanish AIPUs provided data to this study, which means that the results obtained in the different areas studied are quite representative of the reality of the situation of these units in Spain. This is also supported by the participation of AIPUs from all Regions, as well as a high level of participation of AIPUs from public centers, both general and specialized hospitals, and the participation of hospitals with different numbers of beds.

These findings will allow us to present the ways in which AIPUs in Spain function and to examine the structural and functional differences among the different Spanish Regions (and could also allow comparison with other countries). Analysis of these results will allow areas of improvement to be identified for which possible actions may be designed in response and whose implementation will have a positive impact on the working of the AIPU. At the same time, it will also enable us to lay down the foundations for a continuous monitoring system that will allow the measurement of said impact. Accordingly, we hope that our findings will enable us to draw conclusions to help us achieve excellence in the planning of health care for mental health patients and their families and friends.

**Acknowledgments**

This work was possible thanks to the invaluable collaboration of the professionals from the 117 AIPUs that have collaborated, contributing data to the comparative analysis of units. We especially want to thank you for your time and dedication to the project. This project had the support of the Otsuka-Lundbeck Alliance, consisting of companies that fund this vision and approach to mental health.

**Conflicts of Interest**

Author RR-J has been a consultant for, spoken in activities of, or received grants from Instituto de Salud Carlos III, Fondo de Investigación Sanitaria (FIS), Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid Regional Government (S2010/BMD-2422 AGES; S2017/BMD-3740), JanssenCilag, Lundbeck, Otsuka, Pfizer, Ferrer, Juste, Takeda, Exeltis, Angelini, and Casen-Recordati. Author MF-M is involved in clinical trials and research projects funded by Janssen, Otsuka, Lundbeck, Roche, Acadia, Pfizer, Exeltis, and Servier. None of the other authors have any conflicts of interest.
References


Abbreviations

AIPU: acute inpatient psychiatric unit

©Roberto Rodriguez-Jimenez, Iluminada Corripio, Ricardo Campos, Mario Páramo, Manuel Franco-Martin, Estefanía Segura, Sergio González, José Martínez-Raga. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 07.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

Epidemiology of Malaria in East Nusa Tenggara Province in Indonesia: Protocol for a Cross-sectional Study

Robertus Dole Guntur¹,², SSi, MMathSc; Jonathan Kingsley³,⁴, BScAPP, BAppHSc(Hons), MHort, PhD, GradCertTeachLearn(HEd); Fakir M Amirul Islam¹, BSc (Hons), MSc, MApp Science, PhD, GradCertTeachLearn(HEd)

¹Department of Health Science and Biostatistics, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia
²Department of Mathematics, Faculty of Science and Engineering, Nusa Cendana University, Kupang NTT, Indonesia
³Department of Health and Medical Sciences, School of Health Sciences, Swinburne University of Technology, Melbourne, Australia
⁴Centre of Urban Transitions, Swinburne University of Technology, Hawthorn, Melbourne, Australia

Corresponding Author:
Robertus Dole Guntur, SSi, MMathSc
Department of Health Science and Biostatistics
Swinburne University of Technology
John St
Hawthorn
Melbourne, 3122
Australia
Phone: 61 451361478
Email: rguntur@swin.edu.au

Abstract

Background: Malaria is a global pandemic that results in approximately 228 million cases globally; 3.5% of these cases are in Southeast Asian countries, including Indonesia. Following the World Health Organization (WHO) initiative, Indonesia is in the process of achieving malaria-free zone status by 2030. However, the eastern part of Indonesia, including the East Nusa Tenggara Province (ENTP), still has a disproportionately high rate of malaria.

Objective: The aims of this cross-sectional study are to determine the awareness and knowledge, attitude, and practice toward various aspects of malaria among rural adults and their associated factors, including sociodemographic factors and ethnicities; assess the gap between coverage of, access to, and use of long-lasting insecticide-treated nets (LLINs) among the households; estimate the prevalence of and factors associated with malaria in rural adults; and develop a risk prediction model for malaria.

Methods: A multistage cluster sampling procedure with a systematic random sampling procedure at cluster level 4 was applied to recruit 1503 adults aged 18 years or older from the ENTP. Each participant participated in a face-to-face interview to assess their awareness and knowledge, attitude, and practice toward aspects of malaria, practices of sleeping under LLINs, and history of malaria. Information on sociodemographic, environmental, and lifestyle factors was also documented. The proportion of knowledge, attitude, and practice toward aspects of malaria and their variations across different sociodemographic and ethnic groups will be analyzed using descriptive statistics and chi-square tests. Coverage and access to LLINs will be evaluated based on the WHO recommendations. Malaria risk factors will be analyzed using logistic regression. Multilevel logistic regression will be applied to estimate the risk score for malaria.

Results: Of the total participants, 99.46% (1495/1503) of rural adults from 49 villages in the ENTP participated in a face-to-face interview from October to December 2019. The study results are expected to be published in peer-reviewed journals.

Conclusions: The best malaria risk prediction model will be developed in this study. In this protocol, we developed a methodology to provide new evidence to guide health policy in supporting the ENTP government’s expectation to achieve the malaria-free rating by 2030.

International Registered Report Identifier (IRRID): DERR1-10.2196/23545

(JMIR Res Protoc 2021;10(4):e23545) doi:10.2196/23545

KEYWORDS
malaria; rural population; awareness; risk factors; health policy; World Health Organization
**Introduction**

**Background**

Malaria is a major global health problem, with an estimated 3.9 billion people living at risk of malaria infection [1]. In 2018, the World Health Organization (WHO) reported 228 million cases, 35.5% of which were from Southeast Asian (SEA) countries [1]. The action plan of the region indicates that all countries in the region will be malaria-free zones by 2030 [2]. Two countries, the Maldives and Sri Lanka, have been certified malaria-free areas by the WHO, whereas India and Indonesia are still affected by malaria, contributing 85% and 13% to the total number of malaria cases in the region, respectively [1].

Indonesia is a SEA country, with a total population of approximately 265 million [3]. It has a diverse ethnic composition, with 1340 ethnic groups distributed from Sabang to Merauke [4]. A significant reduction in the transmission of malaria in various provinces in Indonesia has occurred since the country has implemented its national commitment to eliminate malaria. As a result of this commitment, 285 of 514 districts (55.5%) achieved malaria elimination in 2018 [5]. However, none of the districts in Papua, West Papua, Maluku, North Maluku, and the East Nusa Tenggara Province (ENTP), Indonesia, have met the malaria elimination area but are committed to elimination by 2030 [5].

The WHO stated that the entire Indonesian population is at risk of contracting malaria and approximately 6.4% of this population have a high risk [1]. The annual parasite incidence (API) survey in 2018 reported that the national API value was 0.84 per 1000 people and that it varied across the 34 provinces [5]. The highest API value was found in Papua province at 52.99 per 1000 people, and in the ENTP (the focus of the proposed study), the API is 3.42 per 1000 people [5]. Over the past decade, there has been a steady decrease in API at the national level in Indonesia from 1.8 per 1000 in 2009 to 0.84 per 1000 in 2018, with this trend observed in most provinces. Despite the consistent decrease in the API value in the ENTP from 13.7 per 1000 people in 2014 [6] to 3.42 per 1000 people in 2018 [5], the API value is well above the national API. As epidemiological malaria research as well as the knowledge of the significant value of API in the ENTP are limited, this study focuses on malaria in the ENTP.

**Problem Statement and Justification**

This study aims to address existing gaps in data focusing on knowledge, attitude, and practice (KAP) toward aspects of malaria; access to and use of long-lasting insecticide-treated nets (LLINs); and malaria risk factors in the ENTP. Several KAP studies on malaria have been conducted in Indonesia [7-9]. However, most of these studies were conducted in western Indonesia, a categorized malaria-free zone, and most were directed at the subdistrict and village levels. One population-based study of 4050 respondents in North Maluku province indicated that although 93.6% of the population realized that malaria is a dangerous disease, almost all respondents (98%) did not know the main causes of malaria [8]. However, 30% of the respondents in the study were children between the ages of 5 and 9 years, making them unsuitable candidates to measure the level of knowledge of a particular community. Another study focused on the KAP toward aspects of malaria at the province level in Central Java provinces [9]. However, the practice of communities using LLINs was not investigated in that study. Studies in various settings have shown that the practice of communities sleeping under LLINs has reduced the transmission of malaria [10,11], and the WHO has recommended using LLINs as the best method to prevent malaria [12].

The increased coverage of LLINs is a key intervention strategy to reduce malaria in Indonesia [13]. From August to October 2017, the Indonesian government implemented a malaria control acceleration program through the mass distribution of LLINs in 67 districts in 5 provinces in the eastern part of Indonesia (Papua, West Papua, Maluku, North Maluku, and ENTP), and of the 22 districts in the ENTP, 15 received the acceleration program [14]. However, despite a 76% increase in the distribution of LLINs from 2015 to 2017, there has been limited publication about access to and coverage of LLINs at the community level in Indonesia [15]. The universal coverage of LLINs [16] has not yet been investigated in the ENTP. A better understanding of the coverage of these indicators would play an important role in developing strategies for a stronger malaria control program in a country [17].

Several epidemiological studies have been conducted to understand the etiology in Indonesia as part of the global effort to eliminate malaria in the country [18-26]. Studies on the social and demographic aspects of malaria have been conducted in Papua province [18,19], Aceh province [26], Maluku province [20], and North Maluku province [21]. However, the effect of the use of LLINs on malaria infection in rural communities was not investigated in any of these studies. Several studies investigated the risk factors for malaria in the ENTP [27,28]. However, either the sample sizes of those studies were too small or the studies were conducted at the subdistrict and village levels. Moreover, although some studies have been conducted at the population level in the ENTP [22-25], they did not evaluate the impact of malaria knowledge, ethnic variations, and coverage of LLINs on the transmission of malaria. Examining the determinant factors of malaria more comprehensively would provide a better understanding of malaria epidemiology and enable experts to identify the important predictors of malaria risk in various environmental settings [29]. A recent study showed that factors associated with self-reported malaria varied between provinces, indicating that the local determinants of malaria risk factors existed at the individual, household, and community levels [20]. Therefore, this cross-sectional study will fill these gaps with the following objectives:

1. Determining awareness and KAP among adults toward various aspects of malaria and their associated factors, including sociodemographic and ethnic groups.
2. Assessing the gap between coverage of and access to and use of LLINs within households.
3. Estimating the prevalence of and factors associated with malaria in rural adults in the ENTP.
4. Developing a risk prediction model for malaria to tailor appropriate interventions.
This study is expected to provide significant findings to comprehensively explain the epidemiology of malaria in the ENTP. The gaps in the knowledge of malaria, the practices of communities using core prevention methods such as LLINs, the practices of malaria treatment–seeking behavior of communities of various ethnicities, and the main malaria risk factors will be identified. These results will help public health policy makers in Indonesia to develop local context-based malaria policies as part of the global effort to achieve a malaria-free zone in Indonesia by 2030. This model can then be implemented in similar socioeconomic settings in other countries.

Methods

Study Population
The ENTP, located in the eastern part of Indonesia, is one of the 34 provinces in the country. The total population is approximately 5.3 million, comprising 2.6 million males and 2.7 million females [30]. This project was conducted in 3 districts—East Sumba, Belu, and East Manggarai—based on the API values of malaria in the region. The East Sumba district has an estimated number of households of about 52,176 [30] and the highest API [5]. The East Manggarai district has an estimated number of households of approximately 55,372 [30] and has the lowest API [5]. Belu district has an estimated number of households of about 46,865 [30] and a moderate API [5]. The 3 districts are shown in Figure 1.

Sample Size and Statistical Power
A cross-sectional study was conducted. The base sample size (n) was calculated using the following formula for a dichotomous outcome in the prevalence study [31]:

\[
 n = \frac{Z^2 \times p \times (1-p)}{d^2}
\]

where \( p \) is the prevalence of malaria in ENTP=1.99% [32], \( Z \) is the confidence level at 95% (standard value of 1.96), \( d \) (relative precision)=0.01125, therefore:
Considering an intraclass correlation coefficient (=0.04) to study malaria prevention methods in Indonesia [33] and a cluster size of n=30 adults per village, the design effect (DEFF) was defined as described by the WHO [34]:

The adjusted sample size after considering DEFF is as follows:

Finally, considering an 85% participation rate (y), the required sample size was calculated, as defined by the WHO [34]:

**Sampling Frame**

A multistage cluster sampling procedure with a systematic random sampling procedure at cluster level 4 was applied for this cross-sectional study. First, out of the 22 districts in the province, 3 were selected based on the malaria API. Second, in each selected district, 3 subdistricts were randomly selected. The number of clusters or villages was selected from each subdistrict based on their relative population sizes. Finally, for each village, a systematic random sampling technique was applied to interview 1 adult per household. Data collection was conducted in 49 villages, with 20 to 40 participants per village proportionate to the population sizes of the villages, as shown in Figure 2.

**Figure 2.** Flow chart for selecting clusters and households in ENTP Indonesia. ENTP: East Nusa Tenggara Province.

**Recruitment Strategy**

First, a cover letter from the chief investigator was sent to the Governor of ENTP to seek approval. Second, the project team sought approval from the head of the East Sumba, Belu, and East Manggarai districts. Once the approval letters at district levels had been received, the research team approached the heads of the subdistricts for their approval. After the research team received written approvals from subdistricts and village leaders, the interviewer started approaching the prospective participants. In the selected households, the research team first approached the heads of households for interviews. In case the household heads, either husband or wife, were absent, any residents above 18 years of age could serve as study participants [35]. As we wanted to investigate malaria knowledge of adults in the ENTP, any potential participants less than 18 years of age were excluded from the study.

**Quality Assurance**

Data collectors, with a health education background, participated in a 1-day intensive training session before commencing household surveys. Intensive training was conducted in Borong,
Belu, and Waingapu cities. The main objective of the training is to improve data collectors’ knowledge of the various aspects of malaria and improve their understanding of the importance of strict adherence to the sampling protocol. Through this intensive training, the interview process will run smoothly and reduce the potential coercion of the participants.

**Questionnaires**

A structured questionnaire was modified from a validated questionnaire [36,37]. Overall, there were 6 main parts of the questionnaire. The first section discusses the demographic information of the participants. We collected information on gender, age, education level, occupation, family size, household income, income of the household head, the main material of the house, ownership of durable assets, access to drinking water, those nearest to the health facilities, and distance to the nearest health facilities. The second part of the questionnaire focused on the general knowledge of malaria. In this part, we asked participants about malaria knowledge, including symptoms of, main causes of, and preventative action to prevent malaria. Next, we collected information on the treatment-seeking behavior of malaria. In this part, we collected information on when and where participants would find the treatment if they or their family members are affected by malaria. Personal protection practices of participants were collected in the fourth section of the questionnaire. In the fifth section, we collected information on self-reported malaria of participants and how they treated their malaria. Finally, at the end of the questionnaire, we obtained data on the demographic information of family members, particularly for those with children aged below 5 years. The practice of sleeping under bed nets for children was also documented. The comprehensive questions for each section are presented in the questionnaire, as shown in Multimedia Appendix 1.

**Outcome Variables**

The outcome variables of the study will be divided into 10 themes. The first outcome variable is the malaria awareness index. This index will be measured by 10 questions related to malaria knowledge on symptoms of, main causes of, prevention actions for, and treatment-seeking behavior for malaria. In each question, 1 mark will be awarded for each correct answer, whereas incorrect answers will be given no points. The total marks for the 10 questions will be evaluated as follows. The level of knowledge with an accurate rate above 80%, 60% to 79%, 1% to 59%, and 0% will be classified as excellent, good, poor, and zero level of awareness, respectively. The rank of excellent and good will be categorized as having malaria awareness, whereas poor and zero levels of awareness will be classified as unaware of malaria [38,39].

The second outcome variable is the predictor of malaria awareness. Malaria awareness of participants will be tabulated based on sociodemographic and environmental factors.

The third outcome variable is the predictor of misconception of the main symptoms and transmission mode of malaria. The main symptom of malaria is fever [40]. All respondents who mentioned other symptoms for the main symptoms of malaria will be classified as misconception. The transmission mode of malaria is mosquito bites [41,42]. All respondents stating other transmission modes of malaria will be classified as misconception. The misconception about the main symptoms and transmission mode of malaria for all participants will be tabulated based on their sociodemographic and environmental factors.

The fourth outcome variable is the gap between knowledge and practice of malaria prevention measures. Knowledge and practice of various malaria prevention measures will be identified for each respondent. The gap between knowledge and practice will be calculated. Finally, the gap between knowledge and practice will be tabulated based on their sociodemographic and environmental factors.

The fifth outcome variable is the malaria treatment–seeking behavior of participants. Appropriate malaria treatment–seeking behavior (AMTSB) is defined as seeking treatment from professional health centers or at health facilities within 24 hours of symptom onset [43]. All respondents stating that they will visit professional health centers within 24 hours were classified as having AMTSB. The AMTSB of participants will be tabulated based on their sociodemographic and environmental factors. Next, the poor understanding of AMTSB is defined as seeking treatment after 24 hours or at nonhealth facilities [44]. All respondents stating that they will visit professional health centers beyond 24 hours or seek treatment at nonhealth facilities will be classified as having a poor understanding of AMTSB. The poor understanding of AMTSB will be tabulated based on their sociodemographic and environmental factors.

The sixth outcome variable is the predictor of bed net ownership and usage among adults. Bed net ownership will be identified for each respondent. The types of bed nets owned are LLINs, non-LLINs, or both of them. This ownership will be classified for each respondent. The use of bed nets and the type of bed net used the previous night (sleeping under the bed net on the previous night before the interview) will be identified for each respondent. Finally, bed net ownership and usage will be tabulated based on their sociodemographic and environmental factors.

The seventh outcome variable is the coverage of LLINs, universal access to LLINs, and the used gap of LLINs. Following the method recommended by the WHO [12,17], 6 indicators will be investigated. They are the proportion of households with at least one LLIN (P1); the proportion of households with at least one usable LLIN (P2); the proportion of households with access to LLINs (P3); the proportion of households with access to LLINs if any LLINs are present (P4); the proportion of households using LLINs the previous night (P5); and the proportion of households that used LLINs the previous night if accessed (P6). Finally, the access gap will be defined as 1-P3, and the use gap will be defined as 1-P6 [17].

The eighth outcome variables are bed net usage and associated factors among children aged below 5 years in the ENTP. The proportion of children aged below 5 years who slept under an LLIN the previous night (P7) is one of the important indicators recommended by the WHO [16]. The use of bed nets and the type of bed net used the previous night (LLINs and non-LLINs) will be identified for each respondent having children aged...
below 5 years. P7 will be tabulated based on the sociodemographic and environmental factors of the parents.

The ninth outcome variable is malaria prevalence and its associated factors. The prevalence of malaria based on self-reported malaria among respondents will be reported. An adult in a household was asked whether they had been diagnosed with positive laboratory-confirmed malaria by local health providers or physicians in the past 12 months. The accuracy of the response was validated by asking the respondent a supplementary question about the symptoms of their malaria events. Approximately 20% of the responses were also validated by contacting their health care providers.

The 10th outcome variable is the prediction of the village with the highest risk of malaria. All villages will be ranked based on the malaria awareness of the community, level of misconception of main symptoms and transmission mode of malaria, gap between knowledge and practice of malaria prevention, level of treatment-seeking behavior, and universal access to bed nets.

Statistical Analysis

For the first outcome, participants' sociodemographic characteristics, including gender, age group, education level, and socioeconomic status (SES), will be reported using descriptive statistics. A chi-square test will be applied to evaluate the association of basic malaria understanding, basic malaria knowledge, the level of malaria knowledge, and the level of malaria awareness among 3 types of malaria endemic settings (MESs). A P value <.05 will be considered statistically significant. SPSS version 27 (IBM Corporation) will be used for analyses.

For the second outcome, participants' sociodemographic and environmental characteristics, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities will be reported using descriptive statistics. A chi-square test will be applied to evaluate the association between malaria awareness and the sociodemographic and environmental characteristics of respondents. Bivariate logistic regression will be performed to explore the association between predictors and malaria awareness of participants. Potential explanatory predictors available with statistically significant differences (P<.05) will finally remain in this model.

For the third outcome, sociodemographic and environmental characteristics of participants hearing malaria terms, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. Perception of participants on malaria symptoms, main symptoms, transmission mode of malaria, misconception of main symptoms, and transmission mode of malaria will be reported in percentage. A chi-square test will be applied to evaluate the association between the misconception and sociodemographic and environmental characteristics of respondents. Bivariate logistic regression will be performed to explore the association between the predictors and the misconception of participants. Potential explanatory predictors available with statistically significant differences (P<.05) will finally be retained in this model.

For the fourth outcome, sociodemographic and environmental characteristics of participants hearing malaria terms, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. The perception of participants on knowledge and practice of the 13 malaria prevention methods will be presented in the proportion. The gap between knowledge and practice in the 13 malaria prevention methods will be presented in the proportion. The level of malaria prevention measures knowledge and practice of sleeping under any bed net to prevent malaria will be tabulated by different sociodemographic and environmental characteristics. A chi-square test will be applied to evaluate the association between knowledge and practice of sleeping under any bed net to prevent malaria and the sociodemographic and environmental characteristics of respondents. Bivariate logistic regression will be performed to explore the association between predictors and knowledge and practice of sleeping under any bed net to prevent malaria. Potential explanatory predictors available with statistically significant differences (P<.05) will finally be retained in this model.

For the fifth outcome, the sociodemographic and environmental characteristics of participants, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. Perception of finding treatment if respondents or their family members have any symptoms of malaria will be presented in the proportion. The proportion of respondents seeking malaria treatment after 24 hours and at nonhealth facilities and having a poor understanding of AMTSB will be tabulated by sociodemographic and environmental characteristics of participants. Bivariate logistic regression will be performed to explore the association between predictors and the perception of respondents on seeking malaria treatment after 24 hours and at nonhealth facilities and having a poor understanding of AMTSB. Potential explanatory predictors available with statistically significant differences (P<.05) will finally be retained in this model.

For the sixth outcome, the sociodemographic and environmental characteristics of participants, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. The ownership and use of the bed net of participants will be presented in the proportion. The ownership of any bed net, LLINs, and non-LLINs will be tabulated by the different sociodemographic characteristics of respondents. Bivariate logistic regression will be performed to explore the association
between predictors and the ownership of any bed net, LLINs, and non-LLINs. The use of any bed net, LLINs, and non-LLINs will be tabulated by different sociodemographic and environmental characteristics of respondents. Bivariate logistic regression will be performed to explore the association between predictors and the use of any bed net, LLINs, and non-LLINs. Potential explanatory predictors available with statistically significant differences ($P<.05$) will finally be retained in this model.

For the seventh outcome, the sociodemographic and environmental characteristics of participants, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. Applying the method provided by the Roll Back Malaria Monitoring and Evaluation Reference Group [16], the gap between the coverage of and access to and use of LLINs will be evaluated. First, the 6 main indicators will be calculated.

The first indicator is the proportion of households with at least one LLIN (P1). The numerator consists of all households with at least one (or 2) LLINs, and the denominator is the total number of sampled households. The second indicator is the proportion of households with at least one usable LLIN (P2). The numerator consists of all households with at least one usable (no visible rift on the net) LLIN, and the denominator is the total number of sampled households.

The third indicator is the proportion of households with access to LLINs (P3). This indicator is defined as households with at least one LLIN for every 2 people and is called sufficient access. The numerator contains all households with at least one LLIN for every 2 people, whereas the denominator is the total number of sampled households. The fourth indicator is the proportion of households with access to LLINs, if any LLINs are present (P4). The numerator contains all households with at least one LLIN for every 2 people, whereas the denominator is the total number of sampled households with at least one LLIN.

The fifth indicator is the proportion of households using LLINs the previous night (P5). The numerator contains all households whose members slept under LLINs the previous night, and the denominator is the total number of households in the sample. The sixth indicator is the proportion of households that used LLINs the previous night if accessed (P6). The numerator contains all households whose members slept under LLINs the previous night, and the denominator is the total number of households with sufficient access to LLINs.

The coverage indicators will be investigated by P1 and P2. The access indicators will be identified by P3 and P4. The use indicators will be investigated by P5 and P6. Households not having at least one LLIN for every 2 people (1-P3) are defined as having insufficient access to LLINs or having an access gap. Households that did not have access to LLINs despite possessing LLINs, which is 1-P4, are defined as intrahousehold net gaps. A household that did not use LLINs the previous night despite having access, which is 1-P6, is defined as the used gap of LLINs.

Descriptive analysis will be conducted to show the distribution of different characteristics of the respondents based on these 6 indicators. To evaluate the association between sociodemographic and environmental characteristics and LLINs’ coverage, access, and use, the chi-square test will be applied. For the eighth outcome, the sociodemographic and environmental characteristics of participants, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. The use of bed nets (P7a), LLINs (P7b), and non-LLINs (P7c) by children aged below 5 years will be calculated in the form of proportion. P7a, P7b, and P7c will be tabulated based on the sociodemographic and environmental factors of the parents. Bivariate logistic regression will be performed to explore the association between predictors and the use of any bed net, LLINs, and non-LLINs by children aged below 5 years. Potential explanatory predictors available with statistically significant differences ($P<.05$) will finally be retained in this model.

For the ninth outcome, the sociodemographic and environmental characteristics of participants, including gender, age group, education level, ethnicity, SES, family size, health facilities close to their house, and distance to the nearest health facilities, will be reported using descriptive statistics. Self-reported malaria of participants will be tabulated based on the 4 aspects of malaria, including malaria awareness, level of malaria knowledge, level of misconception of main symptoms and transmission mode of malaria, knowledge and practice of malaria prevention methods, level of treatment-seeking behavior, bed net ownership and usage, and sufficient access to LLINs. All associations between covariates and self-reported malaria will first be examined using bivariate logistic regression. All significant variables with $P$ values of Wald test $\leq 0.01$ will be considered statistically significant and will be included in the final multivariable logistic regression. The risk prediction model for malaria will be developed based on the significant factors associated with the prevalence of malaria. The prediction value will be assigned based on the severity of the factors, either with binary or multicategory factors.

For the 10th outcome, a risk prediction model for malaria will be developed based on the significant factors associated with the prevalence of malaria. Prediction values will be assigned based on the severity of the factors by applying multilevel modeling. The village variables will be approximated by aggregating individual variables at the village level. The village-level variables will be the location of the village, average distance to health facilities, accessibility to the village, coverage of LLINs, level of malaria knowledge, average education level of the community, and wealth quintile. The logistic regression framework will be applied for modelling the unadjusted, adjusted, and final analyses. Once the best model is identified, a malaria risk scoring system will be developed and the allocation of points for each variable will be based on the magnitude of its regression coefficients [45]. The sum of points for each variable will be used as an approximation to rank the village from the highest to the lowest risk of malaria in the future [45].
Ethics and Dissemination
The Declaration of Helsinki was adhered to ensure that the rights, integrity, and confidentiality of the respondents are strictly protected. All respondents signed consent forms before being interviewed. For this study, we received human ethics approval from the Swinburne University of Technology Human Ethics Committee (Ethics ID: 20191428-1490) and from the Health Research Ethics Committee of the National Institute of Health Research and Development of the Indonesian Ministry of Health (Ethics ID: LB.02.01/2/KE.418/2019). The results of this cross-sectional study will be presented at international conferences and published in peer-reviewed journals.

Results
The time frame of the project is presented in Table 1.
Primary data were collected from October to December 2019. From a total of participants, 99.46% (1495/1503) of rural adults from 49 villages in the ENTP participated in a face-to-face interview. The project is in progress to draft papers that will be published in peer-reviewed journals.

Table 1. Timeline of research phases.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire and question guide development</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ethics approval</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training of field workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household surveys</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data entering</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Data analysis</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Manuscript preparation and submission</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Discussion
Principal Findings
Our study endeavored to explore the contribution of differential risk factors for malaria in rural ENTP. We consider KAP as an important societal risk factor given that contemporary studies in South Asia revealed that low levels of general knowledge on transmission and prevention of malaria in the community in the region had been identified both in the general public and community health practitioners [46]. A study in Myanmar [47] and Nepal [48] highlighted that there is a misconception in the malaria transmission mode in the community and that poor malaria knowledge leads to poor health treatment-seeking behavior for malaria [47]. This study will provide a unique opportunity to identify the gaps in KAP of various aspects of malaria in rural communities of the ENTP. These findings may be published in prestigious journals allowing other malaria experts to compare and contrast the malaria KAP of communities from various settings, including from the ENTP.

This will be the first population-based study revealing KAP toward aspects of malaria in rural adults from 3 different MES in the ENTP. We allocated the balance sample from 3 different MES, allowing us to compare the characteristics of rural adults of their malaria knowledge, practice of malaria prevention method, and treatment-seeking behavior. Furthermore, the study participants selected were solely those living in rural areas, so the true characteristics of the rural population in various aspects of malaria would be revealed through this study. A good understanding of the malaria KAP of the rural population is critical considering that most malaria cases globally were from rural populations [15,49,50]. In the Indonesian context, more than half of the malaria cases were contributed by rural communities [51].

The instruments used to collect the data were modified from a validated questionnaire with some modifications to capture the ownership and use of LLINs both in adults and children aged below 5 years in the ENTP. The assessment of the gap between the coverage of and access to and use of LLINs in the ENTP will be conducted based on the standardized indicators recommended by the WHO [16], allowing the findings to be compared with other national and international studies. Furthermore, the universal access of LLINs and their impact on malaria prevalence in the region could be revealed throughout this study. These findings will provide the best reference for malaria control programs in the region to support the Indonesian government’s expectation of achieving a malaria-free rating by 2030.

Despite the several benefits of this study, there are some limitations. First, adults were asked about their medical histories in the last year, specifically whether they or their children aged below 5 years had contracted malaria. This may reflect recall bias. Next, adult self-reported data will be used to estimate malaria prevalence and investigate the practices of prevention activities and the health and treatment-seeking behavior of communities. This might lead to the introduction of courtesy bias. However, the instruments used to obtain the data applied validated and reliable questionnaires that have been applied in various settings. The instruments were then administered by experienced local health workers.

Conclusions
This research will be expected to provide significant findings to comprehensively explain the epidemiology of malaria in the ENTP. The gap in malaria knowledge, practice of communities in using core prevention methods such as LLINs, practice of malaria treatment–seeking behavior of communities with various ethnicities, and the main malaria risk factors will be recognized.

https://www.researchprotocols.org/2021/4/e23545

(page number not for citation purposes)
These results could help stakeholders in the region to develop malaria policy based on the local context to support the global effort to become a malaria-free zone by 2030.

Acknowledgments
This work was supported by the Australia Awards Scholarship (ST000TBK6) and the Faculty of Health, Arts and Design of the Swinburne University Technology. The funders had no role in the designing of the study, data collection, analysis, or interpretation of data or writing the paper.

Authors' Contributions
RG developed the study protocol, prepared the data collection tools, conducted primary data collection, and analysis plan, and wrote the draft paper. FI and JK supervised the study, reviewed the paper, and provided substantial input. All authors have approved the manuscript for submission.

Conflicts of Interest
None declared.

Multimedia Appendix 1
The questionnaire for primary data collection.
[PDF File (Adobe PDF File), 364 KB - resprot_v10i4e23545_app1.pdf ]

Multimedia Appendix 2
Ethics Approval of Swinburne University of Technology.
[PDF File (Adobe PDF File), 151 KB - resprot_v10i4e23545_app2.pdf ]

Multimedia Appendix 3
Ethics Approval of Indonesia Health Ministry.
[PDF File (Adobe PDF File), 1114 KB - resprot_v10i4e23545_app3.pdf ]

Multimedia Appendix 4
Original peer-review reports from the Swindburne University of Technology.
[PDF File (Adobe PDF File), 2609 KB - resprot_v10i4e23545_app4.pdf ]

References


49. Malaria Situation in South East Asia Region Malaria Unit, Department of Communicable Diseases, WHO SEARO Internet. Ortega LI. 2011. URL: https://www.mmv.org/sites/default/files/content/event_presentation/files/Ortega_Southeast_Asia.pdf [accessed 2021-08-08]


**Abbreviations**

**AMTSB:** appropriate malaria treatment–seeking behavior  
**API:** annual parasite incidence  
**DEFF:** design effect
Protocol

Rehabilitation Needs, Service Provision, and Costs in the First Year Following Traumatic Injuries: Protocol for a Prospective Cohort Study

Helene Lundgaard Soberg1,2*, PhD; Håkon Øgreid Moksnes3,4*, MD; Audny Anke5,6, MD, PhD; Olav Røise7,8, MD, PhD; Cecilie Røe3,8, MD, PhD; Eline Aas9, PhD; Unni Sweeney2,3, PhD; Christine Gaarder8,10, MD, PhD; Pål Aksel Nøes8,10, MD, PhD; Eirik Helseth8,11, MD, PhD; Hilde Margrete Dahl8,12, MD, PhD; Frank Becker8,13, MD, PhD; Marianne Løvstad13,14*, PhD; Kristian Bartnes6,15, MD, PhD; Christoph Schäfer5,6, MD; Mari S Rasmussen1,4, MSc; Paul Perrin16, PhD; Juan Lu17, MD, PhD; Torger Hellstrøm1, MD, PhD; Nada Andelic1,4*, MD, PhD

1Department of Physical Medicine and Rehabilitation, Oslo University Hospital HF, Oslo, Norway
2Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway
3Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway
4Institute of Health and Society, Research Centre for Habilitation and Rehabilitation Models & Services (CHARM), Faculty of Medicine, University of Oslo, Oslo, Norway
5Department of Rehabilitation, University Hospital of North Norway, Tromsø, Norway
6Department of Cardiothoracic and Vascular Surgery, University Hospital of North Norway, Tromsø, Norway
7Division of Orthopaedic Surgery, Oslo University Hospital, Oslo, Norway
8Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
9Department of Health Management and Health Economics, Institute for Health and Society, University of Oslo, Oslo, Norway
10Department of Traumatology, Oslo University Hospital, Oslo, Norway
11Department of Neurosurgery, Oslo University Hospital, Oslo, Norway
12Department of Child Neurology, Oslo University Hospital, Oslo, Norway
13Sunnaas Rehabilitation Hospital, Nesoddtangen, Norway
14Department of Psychology, University of Oslo, Oslo, Norway
15Institute of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway
16Departments of Psychology and Physical Medicine and Rehabilitation, Virginia Commonwealth University, Richmond, VA, United States
17Division of Epidemiology, Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, VA, United States

* these authors contributed equally

Corresponding Author:
Helene Lundgaard Soberg, PhD
Department of Physical Medicine and Rehabilitation
Oslo University Hospital HF
PO Box 4956 Nydalen
Oslo, N-0424
Norway
Phone: 47 951009637
Email: h.l.soberg@medisin.uio.no

Abstract

Background: Traumatic injuries, defined as physical injuries with sudden onset, are a major public health problem worldwide. There is a paucity of knowledge regarding rehabilitation needs and service provision for patients with moderate and major trauma, even if rehabilitation research on a spectrum of specific injuries is available.

Objective: This study aims to describe the prevalence of rehabilitation needs, the provided services, and functional outcomes across all age groups, levels of injury severity, and geographical regions in the first year after trauma. Direct and indirect costs of rehabilitation provision will also be assessed. The overarching aim is to better understand where to target future efforts.
Methods: This is a population-based prospective follow-up study. It encompasses patients of all ages with moderate and severe acute traumatic injury (New Injury Severity Score >9) admitted to the regional trauma centers in southeastern and northern Norway over a 1-year period (2020). Sociodemographic and injury data will be collected. Upon hospital discharge, rehabilitation physicians estimate rehabilitation needs. Rehabilitation needs are assessed by the Rehabilitation Complexity Scale Extended–Trauma (RCS E–Trauma; specialized inpatient rehabilitation), Needs and Provision Complexity Scale (NPCS; community-based rehabilitation and health care service delivery), and Family Needs Questionnaire–Pediatric Version (FNQ-P). Patients, family caregivers, or both will complete questionnaires at 6- and 12-month follow-ups, which are supplemented by telephone interviews. Data on functioning and disability, mental health, health-related quality of life measured by the EuroQol Questionnaire (EQ-5D), and needs and provision of rehabilitation and health care services are collected by validated outcome measures. Unmet needs are represented by the discrepancies between the estimates of the RCS E–Trauma and NPCS at the time of a patient’s discharge and the rehabilitation services the patient has actually received. Formal service provision (including admission to inpatient- or outpatient-based rehabilitation), informal care, and associated costs will be collected.

Results: The project was funded in December 2018 and approved by the Regional Committee for Medical and Health Research Ethics in October 2019. Inclusion of patients began at Oslo University Hospital on January 1, 2020, and at the University Hospital of North Norway on February 1, 2020. As of February 2021, we have enrolled 612 patients, and for 286 patients the 6-month follow-up has been completed. Papers will be drafted for publication throughout 2021 and 2022.

Conclusions: This study will improve our understanding of existing service provision, the gaps between needs and services, and the associated costs for treating patients with moderate and major trauma. This may guide the improvement of rehabilitation and health care resource planning and allocation.

International Registered Report Identifier (IRRID): DERR1-10.2196/25980

(JMIR Res Protoc 2021;10(4):e25980) doi:10.2196/25980

KEYWORDS

trauma; rehabilitation needs; health services; costs; rehabilitation; traumatic injury; injury

Introduction

Traumatic injuries, defined as physical injuries with sudden onset and a severity level requiring immediate medical attention, are a major public health problem worldwide [1]. The most common injuries occur in the extremities (38.3%), head/brain (35.4%), thorax (29.5%), and spine (24.3%) [2]. These injuries are the leading cause of problems in physical, cognitive, emotional, behavioral, and psychosocial functioning, interfering with daily life, work, and health-related quality of life (HRQL) [3-6]. Hemorrhage from internal organs usually does not cause major long-term disabilities. Still, hemorrhage is an important contributor to early deaths and serious complications after trauma including systemic inflammatory response syndrome, multi-organ failure, and acute respiratory distress syndrome [7]. Survivors of these injuries are patients who are catabolic over time, with massive muscle loss, and will have high rehabilitation needs.

While acute trauma care is considered to be of high quality in Norway, rehabilitation services remain fragmented [8,9]. Gabbe et al [4] reported that 80% of major trauma patients suffered problems affecting their daily life 6 months post injury, whereas Soberg et al [5] found that at 2 years post injury, self-reported health and functioning and work participation were significantly lower than in the general population [10,11].

Pediatric trauma constitutes an important subset of traumatic injuries, with lower mortality rates than injuries in adults but a high risk of lifelong functional impairment and disability [12]. Gabbe et al [12] reported that severely injured children showed ongoing disability and reduced HRQL 12 months after injury. Younger children with severe traumatic brain injury (TBI) are at the highest risk of long-term behavioral and learning disabilities and of having unmet or unrecognized health care needs during the first year after injury [13-15].

Trauma rehabilitation should be provided as a set of coordinated, multidisciplinary services tailored to the patient’s needs. Personal factors such as resilience, personality, chronic pain, and access to social support are powerful predictors of outcome [16]; however, the severity of the injury is not a potent predictor of psychological outcomes [17,18]. Thus, these are factors that can affect individuals’ perceived rehabilitation needs and should be taken into account in research on the provision of rehabilitation services.

“A need for rehabilitation” refers to any needs that can be met with rehabilitation management, interventions, and services in the acute, subacute, and postacute phases of an injury. Such needs relate both to specialized inpatient and outpatient services and those provided in community-based settings [19]. A discrepancy between general rehabilitation provision and the need for rehabilitation services has been documented internationally [19]. The extent of service provision may vary across different municipalities and across regions, yet no study has been conducted in Norway to explore how geographical differences in rehabilitation services influence rehabilitation pathways, functioning, and health outcomes of patients with moderate and major trauma [20]. Internationally, research diverges regarding regional differences between high- and low-density population areas. In a Canadian study on multiple trauma survivors, no regional differences were found, but numerous barriers to rehabilitation services were reported [21]. In a survey of rehabilitation services in Australia, Graham and Cameron [22] found that there was better access to rehabilitation...
services in metropolitan areas. Given the paucity of data on the impact of geography on postinjury outcomes, there is a need for a comprehensive evidence base on regional variation in recovery and rehabilitation to optimize postacute services [23]. Few studies exist on the cost-effectiveness of rehabilitation services after traumatic injuries [24]. Patients with complex disabilities after trauma are costly to treat; however, rehabilitation is considered a worthwhile societal investment in the trauma survivors’ regaining of independence and HRQL [25,26]. By applying the Needs and Provision Complexity Scale (NPCS) and a cost assessment algorithm, a provision shortfall of community-based rehabilitation services that may also contribute to increased care burden and costs for family caregivers was reported [27,28] Studies from other countries are needed to confirm these findings. Furthermore, to address shortfalls in service provision to trauma patients, longitudinal, population-based data are required. It can also indicate a baseline for evaluating future system-level changes and service development.

The main aims of this study are to explore the rehabilitation needs and provided services for patients after moderate and major trauma and to study changes in needs over the first year. The study will evaluate the implementation of recommendations on early rehabilitation and continuity of care, as stated in the National Trauma Plan in Norway [8], and assess the specialized and community-based rehabilitation provision. The connections between patients’ rehabilitation needs and functional levels will be explored, and geographical variation in service provision, patient outcomes, and the associated costs will be studied. The main hypotheses are as follows: 1) Underprovision of coordinated rehabilitation services and continuity of care will lead to unmet rehabilitation needs in the trauma patients regardless of type of trauma, demography, postinjury physical and psychological functioning, or regional affiliation. 2) There is geographical variation in patient outcomes for severely injured patients. 3) Rehabilitation is cost-effective; however, there are large societal costs from informal care provided by caregivers.

Methods
Design
This is a multicenter, population-based study with prospective follow-ups at 6 and 12 months post injury.

Study Settings
The regional trauma centers in the southeastern and northern parts of Norway are Oslo University Hospital (OUS) and the University Hospital of North Norway (UNN), respectively. We classify the northern region as more rural due to the long distances to and from hospital and rehabilitation facilities and the low population density. The southeastern region is classified as mainly urban with shorter distances to and from hospital and rehabilitation facilities and higher population density. The prehospital emergency medical services are well organized in both regions [29].

Participants
The group eligible for the study will be approximately 600 patients of all ages with moderate and major trauma, including multiple injuries (extremities, head/brain, thorax, spine, abdomen, face, neck, external/other), who were consecutively admitted to the regional trauma centers in the southeastern and northern parts of Norway over a 1-year period and subsequently discharged.

Inclusion and Exclusion Criteria
All Norwegian residents may be included who were admitted directly to the above hospitals or admitted after transfer from local hospitals within 72 hours after injury, with at least a two-day hospital stay and a New Injury Severity Score (NISS) greater than 9 (ie, moderate to severe injuries). Non-Norwegian residents and deceased patients will be excluded.

Inclusion Procedure
All patients admitted with a trauma alarm will be assessed daily by the participating doctor in the project with formal expertise in scoring the severity of the injuries. Once it is clarified that a patient meets the inclusion criteria, the patient is informed by the research assistant, and informed consent is obtained.

Ethical Considerations, Procedure, and Data Collection
All components of the research project will be conducted according to the ethical guidelines of the Declaration of Helsinki and approved by the Regional Committee for Medical and Health Research Ethics (approval number 31676) and the Norwegian Data Inspectorate (approval number 19/26515). Data will be handled and stored securely on the research servers at OUS and UNN (patient-related data) and the Service for Sensitive Data at the University of Oslo (cost data). All data will be deidentified when sharing them with partners and in analysis and presentations.

Regional research assistants and PhD students will oversee patient recruitment and data collection. The patients will be identified during acute admission to hospital. Trauma patients with NISS>9 are included in both hospitals’ trauma registries and the National Trauma Registry (NTR). It will thus be possible to validate the trauma severity scores of included patients as well as attrition rates in this study. These scores derive from the Abbreviated Injury Scale (AIS), Injury Severity Score (ISS), and NISS.

Information about the study will be presented to patients in written and oral form with informed consent emphasizing their right to withdraw from the project at any time without requiring a reason. The recruited patients (or their proxy or caregiver) must provide this informed consent. For pediatric patients, the information will be given in written and oral form, and forms adapted to the language of the youngest children will be provided. For children and adolescents younger than 18 years, parents must give written consent. Adolescents aged 16-17 years must sign consent for themselves as well. For adult patients with cognitive or communication difficulties, a family member or caregiver will be identified to provide consent (ie, deputy consent) and assist in completing the questionnaires.
Baseline Data Collection and Registration
Collection and registration will be based on information from medical records. The trauma scores will be validated by data registered by certified AIS registrars in the hospitals’ trauma registries and, if necessary, in connection with other available clinical information about the patients’ stays in hospital departments.

Patient Characteristics
Sociodemographic data will include family status (marital status and children), preexisting comorbidity, education or work status, and substance use at the time of injury. Injury-related data include diagnoses (International Statistical Classification of Diseases, Tenth Revision diagnosis codes of S00-T32); cause and type of injury; severity of injury as assessed by the NISS, ISS, and AIS; nonsurgical and surgical treatments; treatment complications; time spent on a ventilator; length of hospital stay; and discharge location.

Injury Severity
The NISS is used to define injury severity in the inclusion phase (scores 9-15 are moderate, 16-24 are severe, and 25-75 are profound). The rationale for the inclusion criterion of an NISS >9 is based on guidance from the National Institute for Health and Care Excellence that such patients should be assessed for rehabilitation needs, and a rehabilitation prescription should be provided for all patients who are deemed to have those rehabilitation needs [30].

Primary Outcomes: Rehabilitation Needs
The Rehabilitation Complexity Scale Extended–Trauma (RCS E–Trauma) is used for identification of needs for specialized inpatient rehabilitation. The RCS E–Trauma is a simple 5-item scale (score range 0-25) that reflects rehabilitation resource requirement for medical, basic care and support, skilled nursing, and equipment needs and therapy inputs [31]. The needs for specialized primary in-hospital rehabilitation (ie, rehabilitation commencing immediately after acute treatment) will be estimated in accordance with clinical judgement by a doctor specialized in rehabilitation medicine at discharge from acute care at the trauma hospital. After 6 months, a medical doctor will use information from a telephone interview with the patient, the caregiver, or both and the medical records to evaluate whether the estimated needs for primary in-hospital rehabilitation were met or not.

The NPCS in its clinician version evaluates needs for community-based rehabilitation and health care service delivery [27]. The NPCS is a pragmatic instrument for measuring both an individual’s needs for rehabilitation and support (“NPCS-Needs”) and the levels of service provided (“NPCS-Gets”) within a given period. An algorithm has been developed to express the impact of met and unmet needs in terms of costs. The NPCS is a 15-item measure that consists of two parts with six subscales and a total score range of 0-50. Part A (NPCS-Needs) is completed by one or more clinicians to evaluate each patient’s needs for health and social care in a given period. Part B (NPCS-Gets) is a mirror image of the same tool, completed at the end of that period, to evaluate the levels of service that have been provided in relation to those needs. Further, the NPCS consists of two main domains. The first is “Health and personal care needs” (score range 0-25), which includes the following subscales: “Health care” (score range 0-6), “Personal care” (score range 0–10), and “Rehabilitation” (score range 0-9). The second is “Social care and support needs” (score range 0-25), which includes the following subscales: “Social and family support” (score range 0-13), “Equipment” (score range 0-3), and “Environment” (score range 0-9) [27]. The expected needs for community-based services during the first 6 months will be estimated in accordance with the clinical judgement of a specialist in rehabilitation medicine upon discharge from the trauma hospital.

The NPCS also includes a self-report version. Patients or carers are asked to report the level of services received within the last 6 months and whether they consider this to be the right amount, too much, or too little (a free-text box is provided for elaboration). The NPCS patient version will be used at 6- and 12-month intervals as an integral part of the telephone interview discussion, measuring the extent to which the self-perceived needs have been met through service provision and informal care. The 6-month report will be used to estimate needs for community-based rehabilitation and health care service delivery for the 6- to 12-month period post discharge. The NPCS will in this study serve as a base for estimation of costs of services received after discharge from the specialist health service.

The Family Needs Questionnaire–Pediatric Version (FNQ-P) [32] will be used to assess the family needs in children across six categories: “Health information,” “Emotional support,” “Community support,” “Instrumental support,” “Professional support,” and “Involvement with care.” Further, formal service provision, including admission to inpatient-, outpatient-, or community-based rehabilitation (as well as informal care) and associated costs will be collected.

Primary and secondary outcomes (ie, measures of functional outcomes, psychological functions, and HRQL) at 6 and 12 months are presented in Table 1.
Table 1. Primary and secondary outcome measures at 6 and 12 months post injury.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Participants</th>
<th>Description</th>
<th>Completed by</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rehabilitation Complexity Scale Extended–Trauma (RCS E–Trauma) [31]</td>
<td>Adult/child</td>
<td>Estimation of specialized inpatient rehabilitation needs</td>
<td>Rehabilitation specialist</td>
</tr>
<tr>
<td>Needs and Provision Complexity Scale (NPCS) A (Needs) and B (Gets) [27]</td>
<td>Adult/child</td>
<td>Estimation of community-based rehabilitation needs and registration of community-based rehabilitation provided (Gets)</td>
<td>Clinician version: rehabilitation specialist Patient version: patients by interview</td>
</tr>
<tr>
<td>Family Needs Questionnaire–Pediatric Version (FNQ-P) [32]</td>
<td>Families of children and youth</td>
<td>A measure for assessing the degree to which the family's needs have been met</td>
<td>Caregivers for children 2-18 years</td>
</tr>
<tr>
<td><strong>Secondary outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glasgow Outcome Scale Extended (GOSE) [33]</td>
<td>Adult/child</td>
<td>A global outcome commonly used in trauma research</td>
<td>Clinician</td>
</tr>
<tr>
<td>Resilience Scale for Adults (RSA) [34]</td>
<td>Adult</td>
<td>A measure of resilience</td>
<td>Patients</td>
</tr>
<tr>
<td>Patient Health Questionnaire 9 (PHQ-9)</td>
<td>Adult</td>
<td>A screening of depression</td>
<td>Patients</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder 7 (GAD-7) [36]</td>
<td>Adult</td>
<td>A screening of anxiety</td>
<td>Patients</td>
</tr>
<tr>
<td>Impact of Event Scale–Revised (IES-R) [37]</td>
<td>Adult</td>
<td>A measure of presence of subjective distress in adults</td>
<td>Patients</td>
</tr>
<tr>
<td>Children’s Revised Impact of Event Scale (CRIES-8) [38]</td>
<td>Child/parent</td>
<td>A measure of subjective distress in pediatric cases</td>
<td>Children from 8 years of age; caregivers</td>
</tr>
<tr>
<td>Strengths and Difficulties Questionnaire (SDQ) [39]</td>
<td>Child</td>
<td>A brief behavioral measure (4-17 years)</td>
<td>Caregivers for children 4-17 years Children 11-17 years</td>
</tr>
<tr>
<td>Return to work/school</td>
<td>Adult/family caregiver/child/parent</td>
<td>Work/school status, return to same role/position, benefits from the labor welfare system at individual and family levels</td>
<td>Clinicians by interview</td>
</tr>
<tr>
<td>WHO(^a) Disability Assessment Schedule 2.0 (WHODAS 2.0) [40]</td>
<td>Adult</td>
<td>A measure of functioning and disability: cognition, mobility, self-care, getting along, life activities and participation</td>
<td>Patients/caregivers</td>
</tr>
<tr>
<td>EuroQol Questionnaire (EQ-5D) [41]</td>
<td>Adult/family caregiver/child/parent</td>
<td>A generic measure of health status (mobility, self-care, usual activities, pain/discomfort, anxiety/depression); health profiles and a weighted total value for HRQL</td>
<td>Patients/caregivers                     Parent proxy for children 0-15 years Children: 8-15 years</td>
</tr>
<tr>
<td>Pediatric Quality of Life Inventory (Peds-QL) 4.0 Generic Core Scales [42]</td>
<td>Child/parent</td>
<td>Patient’s and parent’s perceptions of quality of life</td>
<td>Parents: parent proxy for toddlers (2-4 years), young children (5-7 years), children (8-12 years), and teens (13-18 years) Children/teens: self-report for 5-7 years, 8-12 years, and teens (13-18 years)</td>
</tr>
</tbody>
</table>

\(^a\)WHO: World Health Organization.

All measures exist in Norwegian versions. The estimated time taken for the follow-up interviews and questionnaires at 6 and 12 months is approximately 1.5 to 2.5 hours.

**Follow-up Data Collection**

Questionnaires for patients, their caregivers, or both will be completed at the 6- and 12-month follow-ups by mail or electronic link supplemented by a telephone interview. This will provide data on changes in sociodemographic characteristics, measures of functioning and disability, mental health, and HRQL, as well as rehabilitation, health and social care service needs, and provision of those services (Gets). The discrepancy between the scores of RCS E–Trauma and NPCS assessed at discharge (Needs) and the rehabilitation services received (Gets) during the first 6 months represents unmet needs. Unmet rehabilitation needs at 12 months are the discrepancy between NPCS-Needs as evaluated by patients at 6 months and NPCS-Gets (receiving services 6-12 months after injury) at 12 months.

**Cost Estimation**

We aim to estimate costs of a traumatic injury from a societal perspective, not only as consequences related to medical treatment but also consequences for work and family situations. Health care utilization encompasses in-hospital stays, including rehabilitation and outpatient appointments, as well as home care and contact with community-based rehabilitation services, multidisciplinary rehabilitation, individual rehabilitation plans,
and other health care services. Further, use of social support professionals, vocational and educational rehabilitation, equipment and assistive technology, income and accommodation, informal care provided by family and significant others, and other cost-related variables will be registered. In this study, participants will record the type and amount of services received during the previous 6 months after 6 months and 12 months following discharge from the trauma hospital. The NPCS patient version will be used at 6- and 12-month intervals as an integral part of the telephone interview discussion and serves as one of the bases for estimation of the costs of services (within the subscales “Health care,” “Personal care,” “Rehabilitation,” “Social and family support,” “Equipment,” and “Environment”) received after discharge from the specialist health service. Loss of productivity due to sick leave and absence from work because of treatment, rehabilitation, or follow-up will be estimated together with costs for the family (both informal and intangible costs).

Total rehabilitation, health care, social support, and societal costs across the periods 0-6 and 7-12 months post discharge will be estimated by combining service utilization and national unit costs. The replacement cost method that evaluates informal care time will be based on the cost of paid professionals at the study time as a “shadow price” for informal care [28]. Intangible costs cannot be directly measured in monetary form, but effects such as pain, joy, or physical and psychological limitations will be assessed using the patient’s and family caregiver’s quality of life according to EQ-5D (EuroQol Questionnaire), which will provide an indication of the patient’s and family caregiver’s HRQL after trauma.

By using the NPCS, an algorithm can be applied to estimate the cost of meeting unmet needs for the purpose of integrating care planning [27]. The costing algorithm was developed in the United Kingdom by professionals with particular experience in managing long-term neurological disabilities and will be adjusted for the Norwegian context of this study. Cost of medication is not included in this study as medication does not present a well-defined rehabilitation service.

Power Calculation

According to the estimates from the NTR, 931 and 110 patients with NISS >9 were respectively admitted to the trauma referral hospitals in the southeastern and northern parts of Norway in 2015. The 30-day mortality rate for this group was about 10%. In addition, we expect the dropout rate to be around 30% at the 1-year follow-up, comprising a 20% refusal rate and 10% of patients lost to follow-up. Accounting for mortality and dropouts, this study expects to include approximately 600 participants in the final study analyses. A power analysis was performed for a predictive model with an outcome variable collected at two time points (6 months and 12 months), a predictor variable (for example, with two subgroups such as women and men), and an interaction term between time and the predictor. This type of model would be a common model employed in the proposed study and likely the most power-consuming model given the additional power needed to detect an interaction effect. With 80% power (1-β), and assuming an r=0.50 correlation between the two measurements of the outcome variable, the estimated sample size of 600 participants would generate enough power to detect all large-sized, medium-sized, and small-sized effects with Cohen $\text{d}^2 \geq 0.06$.

Statistics

Descriptive statistics and graphing techniques will be used to describe rehabilitation and health care needs, service use, and functional outcomes over time, as well as to compare these factors across age groups, injury severity levels, and geographical regions. The aggregated data from the NTR will be used to estimate the national prevalence of rehabilitation needs and service use after traumatic injuries, including the total number of trauma patients with moderate and severe injury according to age, gender, and geographical region. Unmet needs for specialized inpatient rehabilitation will be based on the evaluation of RCS E–Trauma compared to received services by the 6-month follow-up. Further, unmet community-based rehabilitation needs will be calculated by subtracting the total NPCS-Get's score from the total NPCS-Needs score.

To establish whether the unmet rehabilitation needs differ across subgroups, longitudinal mixed models will be used wherein the nested variable is time (eg, discrepancies between the estimates of needs at discharge and rehabilitation services the patient reported receiving at 6 and 12 months). The models will enable identification of factors that predict unmet rehabilitation needs and differences in the needs based on key demographic, injury-related, or service-related variables. Predictor variables in these models will be left as continuous wherever possible in order to maximize potential variability in the predictors (eg, age, level of education, injury severity scores). Categorical variables will be dichotomized (eg, employment status, marital status, geographical regions). Injuries can be grouped as follows: TBI, spinal cord injuries, orthopedic injuries including amputations, thorax and abdominal injuries including hemorrhaging.

The mean costs of both formal and informal rehabilitation and health care services used during the first year will be calculated and compared across these variables. In this study we have defined informal rehabilitation as help given by family or friends and quantified by hours per week. Information about this is collected through the telephone interview and the results of the patient version of the NPCS. Costs are typically skewed with heavy right tails due to patients with severe injuries using services to a great extent and incurring high levels of cost; when estimating predictors for total costs, we will apply log-linear and general linear models. Additionally, the NPCS algorithm for calculating the costs of meeting unmet needs will be used for the purpose of integrating care planning.

Data Linkage

For evaluation of completeness and validity of collected data reflecting hospital admissions and service use during the first year after injury, data linkage will be achieved with the Norwegian Patient Register (containing information on diagnoses, length of hospital stays in somatic and rehabilitation units, and costs according to diagnosis-related groups) and with outpatient consultations in the specialist health service.
Additionally, linking will be performed with the registries with costs arising from consultations and other contact with emergency services and various health professionals and medical specialists, and with individual-based health care and rehabilitation use in municipal health care services. Sick leave and other benefits data will be extracted from the Norwegian Labour and Welfare Administration. The applications for extracting these data will be sent to the respective registries.

**Results**

The project was funded in December 2018 and approved by the Regional Committee for Medical and Health Research Ethics in October 2019. Inclusion of patients began at OUS on January 1, 2020, and at the UNN, Tromsø, on February 1, 2020. As of February 2021, we have enrolled 612 patients. The 6-month follow-up has been completed for 286 patients. Papers will be drafted for publication throughout 2021 and 2022.

**Discussion**

This is the first multiregional study in Norway based on data from the NTR on rehabilitation after trauma. We have not found parallel studies on rehabilitation after moderate and major trauma in an international context. This prospective follow-up project is estimated to collect data from over 600 patients and provide innovative insights into rehabilitation needs and costs, as well as intangible costs for patients and family caregivers and their perception of care provided. The study will establish the national prevalence of rehabilitation needs and costs, highlighting patients at risk of unmet rehabilitation needs. Service provision and unmet needs will be evaluated in lieu of outcomes over a broad array of functional domains, thus providing detailed information regarding specific needs.

The results will be of interest across the whole care chain for trauma patients, as there are strong recommendations to plan and initiate the rehabilitation process early in the acute phase. Further, beyond trauma patients, the results will be of beneficial value since this group is representative of patients with long-term disabilities who are expected to have ongoing needs for rehabilitation services and support in general. The results could inform both clinical decision making at the individual level and population-based service planning and delivery in areas capable of improvements.

The project aligns with several national policy documents and white papers and the National Trauma Plan [8,43,44], all of which favor a major strengthening of rehabilitation care in Norway. All of these documents noted a need for reform in this area. They emphasize the need for high-quality and patient-centered care, attentiveness to users’ voices, collaboration between specialist and community health services, integrated and coherent service provision, and long-term support for patients and their caregivers. The documents also express an imperative for individual plans to secure action chains and networks, so that the service provision coalesces as a coherent whole and transitions between responsible services and organizational levels are achieved. Maintaining a user’s perspective implies that an emphasis must be placed on local and flexible solutions that the user can access in his or her everyday life.

The project will inform the World Health Organization global disability action plan 2014-2021 to strengthen and extend rehabilitation provision [45]. The implementation of results could contribute to reducing the burden of injury and improving the lives of people suffering from the consequences of injuries through the development of increasingly individually targeted service delivery. Therefore, the study may highlight both adherence to, and deviation from, the recommendations of the policy documents.

This study also has some limitations that should be discussed. This is an observational study; thus, the causal inferences will be restricted by the design. However, this longitudinal cohort study has a comprehensive and epidemiological approach, comprising registry data and self-reports from patients and caregivers, and the study design provides a higher external validity than an experimental study. Furthermore, as with any longitudinal study, there will be a risk of dropout. To facilitate study adherence and maintain the lowest possible dropout rates, the research team will be trained regarding the follow-up procedures. Baseline data will be available for all participants, allowing comparisons between those who continue in the study and those who drop out. However, since we include patients who are received by the hospitals with trauma alarms, we will not include all patients with injuries with NISS >9, as some with more moderate injuries are not triaged and met by trauma teams.

In conclusion, this study will provide information about rehabilitation needs and service provision for patients who have sustained moderate and major trauma. New knowledge about met and unmet needs will improve the understanding of the existing service provision, the gaps between needs and services, and the associated costs for treating trauma patients, which could guide improvements of rehabilitation and health care resource planning and allocation.

**Acknowledgments**

We would like to acknowledge financial support received from the Southeastern Norway Regional Health Authority and contributions from the Personal Injury Association, Personskadeforbundet, LTN.

**Conflicts of Interest**

None declared.
References


30. NICE guideline [NG40]. 2016. URL: https://www.nice.org.uk/guidance/ng40/chapter/Recommendations-for-research [accessed 2021-02-16]


35. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]


Abbreviations

AIS: Abbreviated Injury Scale
EQ-5D: EuroQol Questionnaire
FNQ-P: Family Needs Questionnaire–Pediatric Version
HRQL: health-related quality of life
ISS: Injury Severity Score
NISS: New Injury Severity Score
NPCS: Needs and Provision Complexity Scale
NTR: National Trauma Registry
OUS: Oslo University Hospital
RCS E–Trauma: Rehabilitation Complexity Scale Extended–Trauma
TBI: traumatic brain injury
UNN: University Hospital of North Norway

Edited by G Eysenbach; submitted 24.11.20; peer-reviewed by B Smith, P Pasquier; comments to author 13.01.21; revised version received 19.02.21; accepted 09.03.21; published 14.04.21.

Please cite as:
Rehabilitation Needs, Service Provision, and Costs in the First Year Following Traumatic Injuries: Protocol for a Prospective Cohort Study
JMIR Res Protoc 2021;10(4):e25980
URL: https://www.researchprotocols.org/2021/4/e25980
doi:10.2196/25980
PMID:33688841

©Helene Lundgaard Soberg, Håkon Øgreid Moksnes, Audny Anke, Olav Røise, Cecilie Røe, Eline Aas, Unni Sveen, Christine Gaarder, Pål Aksel Näss, Eirik Helseth, Hilde Margrete Dahl, Frank Becker, Marianne Lovstad, Kristian Bartnes, Christoph Schäfer, Mari S Rasmussen, Paul Perrin, Juan Lu, Torgeir Hellstrøm, Nada Andelic. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 14.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

The EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) Study: Protocol for the Design, Implementation, and Evaluation of a Community-Based Combination HIV Prevention Intervention for Female Sex Workers in Baltimore, Maryland

Bradley E Silberzahn1, MA; Catherine A Tomko2, MPH; Emily Clouse2, MPH; Katherine Haney2, BA; Sean T Allen2, MPH, PhD; Noya Galai3, PhD; Katherine H A Footer2, JD; Susan G Sherman2, PhD

1Department of Sociology, The University of Texas at Austin, Austin, TX, United States
2Department of Health, Behavior & Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States
3Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Corresponding Author:
Bradley E Silberzahn, MA
Department of Sociology
The University of Texas at Austin
305 E 23rd St, A1700, RLP 3.306
Austin, TX, 78712
United States
Phone: 1 410 614 3518
Email: BradleySilberzahn@utexas.edu

Abstract

Background: Cisgender female sex workers (FSWs) experience high rates of HIV and sexually transmitted infections (STIs), including chlamydia and gonorrhea. Community empowerment–based responses to the risk environment of FSWs have been associated with significant reductions in HIV and STI risk and associated risk behaviors; however, evaluations of US-based interventions targeting FSWs are limited.

Objective: The objective of this study is to describe the design, implementation, and planned evaluation strategy of an ongoing comprehensive community-level intervention in Baltimore City, Maryland, which aims to improve HIV and STI risk and cumulative incidence among FSWs. The two intervention components are the SPARC (Sex Workers Promoting Action, Risk Reduction, and Community Mobilization) drop-in center and the accompanying comprehensive mobile outreach program. The mission of SPARC is to provide low-barrier harm reduction services to FSWs, with a special focus on women who sell sex and use drugs. Services are provided through a harm reduction framework and include reproductive health and sexual health care; medication-assisted treatment; legal aid; counseling; showers, lockers, and laundry; and the distribution of harm reduction tools, including naloxone and sterile drug use supplies (eg, cookers, cotton, syringes, and pipes).

Methods: The SPARC intervention is being evaluated through the EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) study, which consists of a prospective 2-group comparative nonrandomized trial (n=385), a cross-sectional survey (n=100), and in-depth interviews assessing SPARC implementation (n=45). Participants enrolled in the nonrandomized trial completed a survey and HIV and STI testing at 4 intervals (baseline and 6, 12, and 18 months). Participants recruited from predefined areas closest to SPARC comprised the intervention group, and participants from all other areas of Baltimore were included in the control group.

Results: We hypothesize that addressing structural drivers and more immediate medical needs, in combination with peer outreach, will improve the HIV and STI risk environment, leading to community empowerment, and reduce the HIV and STI cumulative incidence and behavioral risks of FSWs. Data collection is ongoing. A baseline description of the cohort is presented.

Conclusions: In the United States, structural interventions aimed at reducing HIV and STIs among FSWs are scarce; to our knowledge, this is the first intervention of its kind in the United States. The results of the EMERALD study can be used to inform the development of future interventions targeting FSWs and other at-risk populations.
Introduction

Background

Cisgender female sex workers (FSWs), defined as cisgender (ie, assigned female at birth and identifying as female) women who exchange sex for money, drugs, or goods, have 14 times the risk of being infected with HIV compared with those who do not exchange sex [1]. Similarly, cisgender FSWs experience sexually transmitted infections (STIs) at disproportionate rates [2-4]. Infectious diseases are often occupational hazards of sex work, with criminalization exacerbating HIV risk behaviors, including unprotected sex with multiple and high-risk sex partners [5,6]. The small body of research on HIV and STIs among FSWs in the United States finds high rates of HIV and STIs, similar to other contexts globally [1,7,8].

The behavioral HIV and STI risk factors of FSWs are positioned in a broader context of enduring sociostructural vulnerabilities (eg, poverty, stigma, violence, and housing instability) that drive chronic substance use, victimization, and poor mental health, particularly in an illicit sex work market [9-11]. These factors are independently and synergistically associated with engagement in sex work and attendant HIV and STI risk [12,13] and are consistently found to be elevated among street-based and unsanctioned venue-based FSWs (eg, exotic dance clubs) compared with brothel-based FSWs [10,14,15]. The illegal nature of sex work heightens these vulnerabilities [10,14,15]. Substance use can also play a complex role in the lives of FSWs, with high rates being reported among samples of FSWs in a variety of settings [16-18]. Substance use can exacerbate engagement in risky sex and serve as a coping mechanism [19,20]. Sexual and physical abuse also heightens vulnerability to HIV and STI infection for FSWs [11,21,22]. Furthermore, FSWs have high rates of mental health morbidities (ie, depression and anxiety), which are risk factors for and underlying determinants of addiction and HIV and STI acquisition [23,24].

In response to these complicated needs and HIV and STI drivers, structural HIV and STI prevention interventions have been developed globally to address social, political, cultural, and economic factors that shape HIV and STI transmission [25-27]. Progress toward the global target of HIV elimination would not be possible without interventions and programs aimed at reducing high-risk behaviors among marginalized groups such as FSWs [28,29]. One of the most renowned and long-standing community-based structural interventions targeting FSWs, Sonagachi, addresses the burden of HIV among FSWs in Calcutta, India, with a holistic approach, including health clinics, banking cooperatives, and social services [30-32]. Using a community empowerment approach, Sonagachi prioritizes FSWs’ involvement in the implementation of the intervention, recognizes sex work as work, and operates from an understanding of the centrality of structural drivers in facilitating and mitigating the health of FSWs.

The Joint United Nations Programme on HIV and AIDS has recognized community empowerment in sex workers as the best practice for over a decade [33]. A recent meta-analysis found that community empowerment–based responses to HIV in FSWs were consistently associated with significant reductions in HIV, gonorrhea, chlamydia, syphilis, and increased condom use [34]. Although robust research exists in international settings, research on US structural interventions for FSWs is undeveloped [13,35,36]. A recent systematic review of US-based HIV and STI prevention intervention research with FSWs found that of 19 interventions, few were rigorously evaluated and none addressed the complex sociostructural context of risk [37].

The need for US-based interventions is underscored by results from our recently completed SAPHIRE (Sex Workers and Police Promoting Health in Risky Environments) study, a prospective observational cohort study of cisgender (n=250) and transgender (n=63) street-based FSWs in Baltimore [38-46]. Cisgender FSWs showed elevated rates of structural vulnerabilities, including housing instability, food insecurity, limited education, and criminal justice involvement, and injection and noninjection drug use, compared with similarly aged peers. Structural vulnerabilities were associated with high baseline prevalence estimates of STIs, including HIV (5%), chlamydia (10%), and gonorrhea (12%) [40], and the incidence of chlamydia and gonorrhea was 13.7% and 18.2%, respectively, over 12 months [43]. Participants also reported frequent interactions with police and a high prevalence of police, intimate partner–, and client-perpetrated violence [38,44].

Building on these results, we aim to develop and evaluate the effectiveness of a community-based combination HIV and STI prevention intervention in Baltimore, Maryland, to include a range of biomedical (eg, HIV and STI testing and counseling, Title X–funded reproductive health, HIV treatment, drug treatment, and primary care referrals), behavioral (eg, HIV and STI risk reduction education), and structural (eg, financial literacy, legal aid, and housing referrals) services provided through SPARC (Sex Workers Promoting Action, Risk Reduction, and Community Mobilization), a drop-in center and comprehensive outreach program. This study uses a structural determinants framework (Figure 1) to understand the complex and multifaceted nature of FSWs’ HIV risk [47]. This framework is informed by fundamental cause theory, which argues that fundamental causes (eg, poverty) have a greater...
impact on health than behavioral risks, and empowerment theory, which links social participation to increased social cohesion [34,48,49]. In this paper, we describe the design of the EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) study, which evaluates the impact of the SPARC intervention. We also detail the SPARC intervention aims, components, and lessons learned during implementation. All study procedures were approved by the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (JHBPSPH IRB).

**Figure 1.** EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) study theoretical framework. GED: general educational development; SPARC: Sex Workers Promoting Action, Risk Reduction, and Community Mobilization; STI: sexually transmitted infection; SW: sex worker.

**Aims and Hypotheses**

The EMERALD study, detailed below, aims to examine (1) the effect of intervention exposure on HIV and STI risk behaviors (eg, drug use and unprotected sex) and HIV and STI cumulative incidence over time in FSWs in the intervention group compared with those in the comparison group, (2) how sociostructural (eg, social cohesion and stigma) and structural vulnerability (eg, financial and housing stability) indicators change and are associated with the biological and behavioral outcomes over time in FSWs in the intervention group compared with those in the comparison group, (3) the role of these indicators as mediators of the intervention effect on study outcomes, and (4) the implementation of the intervention through qualitative (eg, in-depth interviews) and quantitative (eg, assessment of program fit, reach, and cost, and facilitators and barriers to utilization measures).

The SPARC center and associated outreach is the intervention being evaluated by the EMERALD study, which consists of an 18-month longitudinal cohort study comprising intervention (n=224) and control (n=161) participants, a cross-sectional survey to assess program reach (n=100), and in-depth interviews (n=45) to assess SPARC implementation. We hypothesize that addressing structural drivers and more immediate medical needs, in combination with peer outreach, will lead to community empowerment and reduce FSWs’ HIV and STI incidence and behavioral risks.

**Methods**

**Intervention: The SPARC Center and Outreach Program**

The 2 intervention components are the SPARC drop-in center and the accompanying comprehensive outreach program, both of which target FSWs in the intervention area in south and southwest Baltimore (Figure 2). SPARC opened in November 2017, and the outreach program began in earnest the following fall. The mission of SPARC is to provide low-barrier harm
reduction services to at-risk nonmen, with a special focus on women who sell sex and use drugs. SPARC addresses clients’ needs through nonjudgmental, convenient, safe, and nonstigmatizing interactions, reducing the need for outside referrals and increasing the likelihood of continued engagement in care and service utilization. All services are provided through a harm reduction framework, centering the lived experiences of those served, while addressing the structural and socioeconomic constraints that often limit their choices and options. Ultimately, SPARC’s overarching goal is to foster a sense of community and stimulate empowerment by creating a safe physical space in which women can connect and develop social cohesion and a sense of collectivism.

Figure 2. Map of EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) intervention and control areas. Red star=SPARC (Sex Workers Promoting Action, Risk Reduction, and Community Mobilization) center, blue=water, pink=intervention area, and green=control area.

SPARC’s day-to-day operations are bolstered by key collaborations with community providers who offer services at SPARC, enabling these organizations to reach a population they otherwise do not easily access. The Baltimore City Health Department funds 2 nurse practitioners and a medical assistant to run a weekly reproductive health clinic (eg, long-acting reversible contraceptives, Papanicolaou testing, and intrauterine devices) and a twice-weekly sexual health clinic (eg, pre-exposure prophylaxis, HIV and Hepatitis C testing and treatment referrals, and STI testing and treatment). The Behavioral Health Leadership Initiative funds 2 providers to staff a weekly low-threshold medication-assisted treatment clinic at SPARC. Biweekly law clinics are offered through the Legal Aid of Maryland and the Maryland Volunteer Lawyers Network.

SPARC’s reach is expanded through an intensive mobile outreach program, which brings many tangible harm reduction tools and microcounseling to women while they are working and living on the street. The outreach program is deployed throughout south and southwest Baltimore at night and during the day between 3 and 4 times a week, serving 15 to 40 people per shift. Outreach staff work in teams to provide relevant harm reduction education and supplies from the window of the outreach vehicle, through drop-off supply bags, and during community events.

All SPARC and outreach staff receive extensive training as a part of onboarding and continuing education. Training sessions are conducted by the study staff as well as external experts and cover topics such as trauma-informed care, harm reduction principles and practice, interpersonal and systemic violence, safer drug use (eg, safer injection), naloxone administration, gender and sexual identity, and de-escalation and staff safety practices. Understanding that trauma often serves as a direct barrier to care, much of the continuing education at SPARC
centers around trauma-informed care and supporting the complex needs of vulnerable individuals.

To ensure equal treatment of all guests, the SPARC center and outreach staff are blind to EMERALD study participation. Limited guest information is collected at intake, primarily their name and birthdate, sex work and drug use history, and service needs. This information allows the team to track service utilization among all women to inform programming and determine whether the intervention is reaching the target population. SPARC guests are asked to sign a consent form indicating that they agree to allow the research team to access their SPARC data. Unique identifiers will be linked between EMERALD and SPARC to determine the frequency of SPARC center interactions among EMERALD participants at the conclusion of all EMERALD data collection. Owing to the quick pace and casual nature of outreach, no identifying information is collected on outreach shifts.

The SPARC program contributes to the innovation of a service model that incorporates health, social, legal, and basic needs (eg, emergency food, shower, laundry, and space to relax) services in a convenient space, paired with a comprehensive outreach program that brings services, supplies, and information to the community, increasing the likelihood of engagement in care and continued service utilization.

The EMERALD Study

Study Design

The primary component of the EMERALD study, which evaluates the SPARC center and associated outreach, is a prospective 2-group comparative nonrandomized trial. Participants were recruited into the control and intervention arms. The intervention area is the south and southwest portion of Baltimore City because of high concentrations of street-based FSWs and a dearth of tailored FSW services, indicated by pink shading (Figure 2). The control group participants were recruited primarily from the southeast and northwest Baltimore. The SPARC center is located within the intervention area, and outreach is conducted in areas with potential sex work activity throughout the intervention area.

The study diagram (Figure 3) illustrates participant progress through the study stages, sample size, and follow-up schedule. Participants completed a survey, HIV testing, and STI testing for chlamydia and gonorrhea during study visits at baseline and at 6, 12, and 18 months. In-depth interviews began at the conclusion of the EMERALD cohort and are currently being conducted with intervention area participants, SPARC staff, and outreach workers. In addition, 100 cross-sectional surveys are being collected in the intervention area to assess program reach in the community beyond the EMERALD cohort.

![Figure 3. EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) study flow. STI: sexually transmitted infection; VDT: venue-date-time.](https://www.researchprotocols.org/2021/4/e23412)
Cohort Study Target Population and Sample

Cohort participants were FSWs aged ≥18 years who live in Baltimore City; self-reported that they had traded sex for money, goods, or drugs 3 or more times in the past 3 months; and were not currently enrolled in our previous cohort study, the SAPPHIRE study [40]. Although transgender female sex workers (TFSWs) who reported being assigned male at birth were encouraged to visit the SPARC center and engage with peer outreach workers, they were excluded from the EMERALD study cohort because of a lack of comparable recruitment locations between the intervention and control areas.

Participants were assigned to a study arm based on the geographic location in which they were recruited. The initial target sample size was 350 women based on power calculations to observe a change in the intervention compared with the control group over the 3 follow-up visits. The assumptions used for the calculation were (1) n=200 in the intervention group and n=150 in the comparison group, (2) lost-to-follow-up (LTFU) rate of 10% to 20% in the intervention (effective n=160-180) and LTFU of 20% to 25% in the comparison group (effective n of 120-112), (3) intraperson correlation (Rho)=0.2-0.4, (4) varying levels of baseline risk behavior rate, and (5) a 2-sided α=.05. Thus, assuming effective sample sizes of n=160 in the intervention and n=112 in the comparison groups, the power is 83% to detect an odds ratio of 0.50, assuming a 20% prevalence rate of a risk behavior (eg, drug use) in the control group and an intraperson correlation of Rho=0.2. For the cumulative incidence outcome, with an average N contributing to the analysis of 170 in the intervention and 120 in the control groups, and assuming annual incidence rate of 6%, the power is 80% to detect a rate ratio of 0.28.

The enrollment target was increased to 450 midway through data collection after power calculations were reassessed based on actual retention rates within the sample, which were closer to 65%. The LTFU rate was higher than anticipated based on previous studies [41], likely because of high incarceration rates, the mobility of the population, and the length of time between study visits (6 months). Although we aimed for a sample of 450 participants, recruitment was stopped at 385 because of reaching saturation in the recruitment areas and repeatedly encountering participants attempting to reenroll in the study.

Cohort Recruitment

EMERALD recruitment began 2 months before the opening of the SPARC center and took place from September 2017 to January 2019. The recruitment strategy for the EMERALD cohort was informed by SAPPHIRE, an earlier prospective cohort study of FSWs in Baltimore, conducted by this study team, which used targeted sampling [50]. Given that recruitment for SAPPHIRE concluded approximately 9 months before the launch of the EMERALD study, we conducted a series of geospatial analyses of possible sex work activity indicators using publicly available data [51] (eg, prostitution charge data and 911 call center reports of suspected prostitution) to understand possible shifts in the geotemporal distribution of sex work throughout Baltimore City (as compared with where we previously recruited FSWs in SAPPHIRE).

In total, the sampling frame for EMERALD consisted of 10 small geographic areas with high concentrations of sex work activity. Among them, 6 were located in the intervention area and 4 in the control area. Compared with our previous targeted sampling recruitment strategy among FSWs in Baltimore City [50], the time and day components of the sampling frame were selected based on analyses of time signatures associated with relevant secondary data in each location with high concentrations of sex work activity and supplemented with additional insights gleaned from SAPPHIRE. Before launching recruitment, staff conducted windshield tours in each location to confirm the presence of sex work activity.

A recreational vehicle (RV) retrofitted with 2 private interview areas and a bathroom for sample collection was driven to the field locations in the intervention and control areas 3 to 5 days per week depending on the interview volume. Each shift lasted approximately 4 hours at the recruitment site and was staffed by 1 field supervisor and 2 interviewers. A community advisory board made up of current and former sex workers provided input on all data collection strategies and instruments.

Potential participants were discreetly approached by staff who provided a brief description of the study. Due to the sensitive nature of sex work, drug use, and HIV, study staff were trained to refer to the study as a women’s health study when communicating with potential participants or community members. Women who were interested in participating were offered screening in a private area on the study RV. FSWs who were eligible and interested in participating in the study were asked to provide written informed consent for all study procedures using an electronic consent form. Following the consent process, participants completed a standard locator form, which included mobile and home phone numbers, addresses, social media usernames, places frequented by the participant, and contact information of family or friends. This locator information is collected and updated at every visit to increase the likelihood of successful retention for future study visits [41].

After completing the locator form, participants received HIV pretest counseling from trained staff and were tested for HIV using the OraQuick ADVANCE Rapid HIV-1/2 Antibody Test, which detects HIV-1 and HIV-2 antibodies in saliva in 20 minutes. While waiting for the results of their HIV test, participants completed a 45- to 60-minute audio-enhanced computer-assisted self-interviewing (A-CASI) survey, with staff on hand to provide assistance. The use of A-CASI provided uniformity of delivery and afforded greater privacy and confidentiality than an interviewer-administered survey, considering the sensitive nature of some survey questions. The baseline survey assessed sociodemographic characteristics, arrest and prison history, sex work history, frequency and type of police encounters, drug and sexual-risk behaviors, overdose and drug treatment history, general health history, experiences of sexual and physical violence, FSW social cohesion, stigma, empowerment, health service utilization, and a brief screening for posttraumatic stress disorder and depression symptoms (Table 1).
Table 1. Variables reflected in the conceptual framework and collected as part of the EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) study of female sex workers in Baltimore, Maryland.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Example question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcomes and impact</strong></td>
<td></td>
</tr>
<tr>
<td>HIV diagnosis</td>
<td>N/A</td>
</tr>
<tr>
<td>Sexually transmitted infection diagnosis</td>
<td>N/A</td>
</tr>
<tr>
<td>Substance use</td>
<td>Have you (injected, smoked, swallowed, or snorted—asked separately) any of the following drugs?</td>
</tr>
<tr>
<td></td>
<td>• Powdered cocaine</td>
</tr>
<tr>
<td></td>
<td>• Crack cocaine</td>
</tr>
<tr>
<td></td>
<td>• Heroin</td>
</tr>
<tr>
<td></td>
<td>• Fentanyl</td>
</tr>
<tr>
<td></td>
<td>• Buprenorphine or suboxone</td>
</tr>
<tr>
<td></td>
<td>• Prescription pain relievers</td>
</tr>
<tr>
<td></td>
<td>• Sedatives or tranquilizers</td>
</tr>
<tr>
<td></td>
<td>• Stimulants</td>
</tr>
<tr>
<td>Condom use</td>
<td>In the last week, how often did you use condoms when having (vaginal or anal) sex with (clients or intimate partner)?</td>
</tr>
<tr>
<td></td>
<td>• Always</td>
</tr>
<tr>
<td></td>
<td>• Most of the time</td>
</tr>
<tr>
<td></td>
<td>• Sometimes</td>
</tr>
<tr>
<td></td>
<td>• Rarely</td>
</tr>
<tr>
<td></td>
<td>• Never</td>
</tr>
<tr>
<td></td>
<td>• Did not have (vaginal or anal) sex</td>
</tr>
<tr>
<td>Safe drug use practices</td>
<td>In the last 6 months, did you use any of the following items that you know have been used by someone else?</td>
</tr>
<tr>
<td></td>
<td>• Syringes or needles</td>
</tr>
<tr>
<td></td>
<td>• Cookers</td>
</tr>
<tr>
<td></td>
<td>• Cotton</td>
</tr>
<tr>
<td></td>
<td>• None of these</td>
</tr>
<tr>
<td><strong>Confounders</strong></td>
<td></td>
</tr>
<tr>
<td>Housing stability</td>
<td>Have you been homeless in the past 6 months?</td>
</tr>
<tr>
<td></td>
<td>Where do you currently stay?</td>
</tr>
<tr>
<td></td>
<td>• Place that you own</td>
</tr>
<tr>
<td></td>
<td>• Place that you rent</td>
</tr>
<tr>
<td></td>
<td>• Family member’s place</td>
</tr>
<tr>
<td></td>
<td>• Shelter</td>
</tr>
<tr>
<td></td>
<td>• Transitional housing program</td>
</tr>
<tr>
<td></td>
<td>• Hotel or motel</td>
</tr>
<tr>
<td></td>
<td>• Streets, park, car, or abandoned building</td>
</tr>
<tr>
<td>Financial stability</td>
<td>In the past 6 months, did you depend on anyone financially? This includes for food or a place to stay.</td>
</tr>
<tr>
<td></td>
<td>Do you currently owe money to anyone, including a person, company, or a bank?</td>
</tr>
<tr>
<td>Access to physical and mental health services</td>
<td>Was there a time in the past 6 months when you wanted or needed to see a doctor or health care provider (other than addiction services) but could not?</td>
</tr>
<tr>
<td>Social cohesion</td>
<td>You can count on other people who sell sex if you need to borrow money; 4-point Likert scale (strongly disagree to strongly agree)</td>
</tr>
<tr>
<td></td>
<td>You look out for new girls when they start selling sex on the street; 4-point Likert scale (strongly disagree to strongly agree)</td>
</tr>
<tr>
<td>Collective action and social participation</td>
<td></td>
</tr>
</tbody>
</table>
Variable | Example question
--- | ---
Do you participate in any of the following:
- Church
- Clubs
- Cultural activities
- Community organizations

Gender equity
Men’s opinions are more important than women’s in making important decisions in a primary relationship; 4-point Likert scale (strongly disagree to strongly agree)
If a man wants to have sex in a primary relationship and a woman does not, she should have sex to please him; 4-point Likert scale (strongly disagree to strongly agree)

Violent experiences [52]
Thinking about your (clients or intimate partners):
- Have you been hit, punched, slapped, or otherwise physically hurt by them?
- Have they removed a condom during sex after agreeing to use one?

Sex work stigma [53]
There are times you feel ashamed of selling sex; 4-point Likert scale (totally disagree to total agree)
People’s attitudes about selling sex make you feel worse about yourself; 4-point Likert scale (totally disagree to total agree)

Empowerment [54]
I have little control over the things that happen to me; 5-point Likert scale (totally disagree to total agree)
There is little I can do to change many of the important things in my life; 5-point Likert scale (totally disagree to total agree)

On completion of the survey, participants were asked to self-collect vaginal swabs with Aptima vaginal swabs (Hologic Inc) using the private RV bathroom. The samples were sent to the Baltimore City Health Department lab, which conducted nucleic acid amplification testing for gonorrhea and chlamydia. Once samples were collected, the participants received the results of their HIV test and posttest counseling. A 2-week appointment was scheduled to receive the STI test results.

Participants were given a US $70 prepaid Visa debit card as compensation for their time and offered relevant referrals to health and social services. Participants recruited from the intervention area were encouraged to visit SPARC and offered hygiene products such as sanitizer wipes and lip balm with SPARC branding, including the address and phone number of the center. Control area FSWs were offered similar EMERALD-branded items and provided neighborhood-specific referrals if requested but were not encouraged to visit SPARC.

Cohort Retention Methods
Follow-up visits for EMERALD occurred at 6, 12, and 18 months after the initial baseline interview. Participants received US $40 prepaid Visa debit cards for completing each follow-up visit. Participants were eligible to complete each follow-up visit 1 month before their actual interview date until 1 month after, for a total eligibility window of 2 months. Those who failed to complete an interview within the allotted eligibility window were considered lost to follow-up. Participants who missed a visit could still complete their next scheduled visit and were not withdrawn from the study.

Retention protocols were modeled after strategies used in SAPPHIRE [50] and are more fully described in a study by Silberzahn et al [41]. Briefly, retention strategies consisted of the aforementioned locator forms (eg, phone numbers, addresses, and social media accounts); promotional materials (eg, hand sanitizer, wipes, lip balm, and silicon bracelets) branded with the study logo and phone number; monitoring public information databases to determine if participants were incarcerated, had moved, or were deceased; and outreach in the form of phone calls, social media messaging, home visits, and repeated visits to recruitment areas in both the study RV and a study sedan, which was used for follow-up. The study team also relied on tracking teams of 2 study staff who were assigned participants to focus on with 1 month of eligibility remaining.

Building on the strategies used in SAPPHIRE, the EMERALD cohort retention protocols were enhanced using several additional strategies. A password-protected database was developed to monitor follow-up and participant progress during the study visits. Using this database, study management could easily determine current and projected follow-up rates, the number of participants due for a study visit (6, 12, and 18 months), and all previous contact attempts used to locate a participant. Field staff documented how participants were located (eg, phone calls and home visits) and time spent on van shifts and tracking, which allowed study management to...
determine adequate staffing and consider the best follow-up strategies. For example, study staff could monitor the number of difficult-to-reach participants due for follow-up surveys in specific study areas and direct tracking teams to these locations to increase the probability of chance encounters with participants. Additionally, check-in locator visits were conducted in between formal study visits at 3, 9, and 15 months to touch base with participants and inquire about any changes to contact information on file (eg, phone numbers, addresses, family and friends). Finally, although SPARC staff were prohibited from discussing the EMERALD study with clients, participants could obtain the EMERALD study phone number from SPARC center administrative staff at their request and use the SPARC office phone to contact the EMERALD study staff, determine follow-up eligibility, and schedule follow-up interviews.

**In-Depth Interviews and Cross-sectional Surveys**

In-depth interviews and cross-sectional surveys began at the conclusion of the 18-month EMERALD cohort. In total, thirty 45- to 60-minute in-depth interviews will be conducted with intervention cohort members who did (n=15) and did not (n=15) visit the SPARC center, and 15 interviews will be completed with SPARC center staff and outreach workers. The EMERALD cohort participants receive a US $25 prepaid Visa debit card for completing the interview; however, SPARC staff are not paid for participation. Interviews with intervention area cohort participants assess program fit, facilitators, and barriers to access of SPARC center services. Respondents are asked about the convenience of hours of operation, center location, staff, and services. Participants are also asked what they liked or disliked about the program and recommendations for additional services. Interviews with SPARC staff examine facilitators and barriers to the success of SPARC and how to improve services and reach.

We are further assessing the intervention’s reach (eg, peer educator contact) through a cross-sectional survey at 18 months among cisgender women aged ≥18 years (n=100) recruited from the intervention areas. Surveys are interviewer administered; take approximately 10 minutes to complete; and consist of questions regarding sex work history, knowledge, and utilization of SPARC services. Participants receive a US $5 prepaid Visa card for completing the cross-sectional survey.

**Study Monitoring**

The EMERALD study team has several key roles: the principal investigator and co-investigators, study director, field supervisor, field staff, and data manager. Investigators and the study director are responsible for protocol development and the oversight of the study progress. Field staff are responsible for data collection, and the data manager monitors and analyzes all study data to ensure that the data are of the highest quality. EMERALD investigators and the research team meet weekly to obtain feedback from field staff regarding study protocols and to discuss study progress. On the basis of staff feedback, data collection protocols are updated and subsequently sent to the JHBSPH IRB for approval. Any adverse events are reported to the JHBSPH IRB within 10 days, as required. Annual progress reports are also submitted to the study sponsor and to the JHBSPH IRB, documenting study progress. Finally, an independent and external data safety monitoring board comprising individuals with relevant expertise provides oversight and monitors study progress.

**Statistical Analysis**

The principal analytic approach will center on generalized linear models for the individual data, with variance components that reflect the potential correlations among FSWs recruited from the same area and between repeated measurements from the same individual at baseline and follow-up visits [55,56]. The main covariate of interest is the intervention indicator. To address aim 1, we consider illicit drug use and irregular condom use as the primary behavioral endpoints.

The effect of the intervention will be assessed with a random effect hierarchical logistic model, with the intervention indicator, time, and their interaction as the main explanatory variables. To adjust for possible imbalances between the 2 areas, we will adjust for baseline confounding factors, including age, length of sex work, and number of sexual partners. The intervention effect will be assessed by comparing time trends in the intervention group with the time trend in the comparison group, as captured by the interaction term. The model will be expanded to include additional demographic covariates and time-dependent indicators of events such as arrest, homelessness, use of health services, or exposure to other prevention activities that are unrelated to the study. Aim 1 also seeks to evaluate the association between the degree of exposure to the intervention and observed behavioral changes. We will explore models that include as covariates participation or exposure to specific intervention components (ie, receiving medical care, drug treatment referral, and legal aid). On the basis of the literature, an 18-month follow-up should be sufficient to document meaningful changes in study outcomes, given the extent of the intervention [57].

For the primary biological endpoint of HIV and STI incidence, we will initially estimate cumulative incidence rates in both the intervention and comparison participants as a binary outcome and construct confidence intervals, accounting for the correlation structure because of nesting in geographic recruitment areas. The intervention’s effect will be assessed using an unadjusted random effect logistic model, with the intervention indicator as the only explanatory variable. The model will be expanded to include baseline covariates, such as demographic characteristics, history of sex work, and baseline risk behaviors (ie, condom use). A more detailed analysis will employ person-time methods with Poisson regression models to allow for time-updated covariates measured over time. Those lost to follow-up will contribute time-at-risk until the time of their last visit.

Finally, to evaluate the reach of the intervention, we will estimate SPARC involvement of FSWs in the independent cross-sectional survey at the 18-month follow-up period. We will also compare this sample with the FSW cohort at 18 months to assess their comparability or differences regarding demographic and risk profiles using logistic regression with group membership as the outcome variable.

To address aim 2, we will first describe changes over time in various sociostructural (ie, social cohesion and stigma) and structural vulnerability (eg, housing stability) indicators. We
will then contrast these trends between the intervention and comparison groups. Social cohesion, social participation, and collective action will be analyzed as aggregate measures, the form of which will be determined by exploratory analysis. Specific models for formal statistical comparisons will be developed based on exploratory analysis, identifying both the distributional properties of the indicators and the shape of changes over time. To assess the association of these changes with the study’s primary endpoints, we will develop random effects logistic models, where the changes in structural indicators at each visit compared with baseline will be the main explanatory factors, in addition to the intervention indicator. The random effects will reflect the nesting of observations within the recruitment area and person over time.

To address aim 3, we will carry out mediation analyses evaluating the contribution of intermediate endpoints that were targeted by the intervention on the primary study outcomes. These include sociostructural factors (eg, stigma and social cohesion) and behavioral outcomes (illicit drug use and irregular condom use). We will have limited power to perform a definitive mediation analysis, but we plan to run detailed descriptive analyses and develop initial models to evaluate potential mediators that could be later validated in a larger study. The analysis will be based on recent developments in mediation analysis that extend earlier results to outcomes that are binary, such as irregular condom use. Specifically, the mediation formula proposed by Pearl [58] and the general framework for causal mediation analysis by Imai et al [59] will be adopted. Full data analysis and power calculation details are available from the senior author upon reasonable request.

To address aim 4, interviews with staff, FSWs, and peer educators will be entered and managed separately using ATLAS.ti software (Scientific Software Development GmbH). Coding will reduce the data to manageable units of information that cover broad and general categories. Themes that emerge from the data will be analyzed using a grounded theory approach [60,61]. A total of 2 coders will conduct open coding on 3 transcripts to develop initial coding schemes. After discussion and development of a combined draft scheme, 2 additional interviews will be coded to inform the final coding scheme. Discrepancies will be discussed with the interviewers and the primary investigator. Through weekly analysis meetings, a team approach to data analysis will be employed, whereby different researchers provide feedback on emerging interpretations and check emerging categories against the raw data. An audit trail will ensure the trustworthiness of findings, gather input from multiple perspectives, and enhance reliability [62].

Results

Recruitment and Retention
At the conclusion of EMERALD cohort recruitment, 93.9% (596/635) of the women who were approached agreed to be screened for eligibility. Reasons for ineligibility of women included: no history of selling sex (n=43), not selling sex 3 or more times in the past 3 months (n=48), not cisgender (n=2), currently enrolled in the SAPPHIRE study (n=15), currently enrolled in EMERALD (ie, enrolled participant rescreening: n=6), and inability to complete study procedures (eg, too intoxicated to give consent, too tired, and not available for enough time; n=12). A total of 470 women who were screened were eligible, provided written informed consent, and completed all baseline study procedures. After the baseline screening, 85 women were withdrawn from the study, and their data were removed from the data set. The reasons for withdrawal included already enrolled or duplicate (n=57), protocol issues such as incomplete data (n=13), participant chose to withdraw (n=2), or other reasons (n=13). When duplicates were discovered, only the first survey was included in the data set. For a participant to be removed for incomplete data, substantial portions of the survey needed to be incomplete, such as skipping a majority of sections or choosing “refused to answer” or “don’t know” for a majority of survey answers. This resulted in a final analytical sample of 385.

Cohort Participant Characteristics
The final sample of the cohort (n=385) had a mean age of 37 years (SD 9.3 years), 56.6% (218/385) of participants were White, 70.9% (273/385) had attained less than a college education, and approximately two thirds of the sample (257/385, 66.8%) had a recent history of homelessness (Table 2). There were significant differences in racial categories between the intervention and control areas (P=.04). Women reported an average of 13 years (SD 9.5 years) in sex work. The prevalence of current substance use was high: 57.9% (223/385) injected any drug, 80.3% used heroin via any method, and 86.8% (334/385) used powdered or crack cocaine in the past 6 months. Almost half of the sample (174/385, 45.2%) reported a recent condomless sex with clients. HIV prevalence in the full sample was 5.2% (20/385), although only 3 women received new HIV diagnoses from study-related testing. Baseline gonorrhea and chlamydia prevalence were high at 15.8% (59/373) and 18.2% (68/374), respectively. There were no significant differences between the control and intervention groups in the key study outcomes at baseline.
Table 2. Baseline characteristics of a sample of female sex workers in Baltimore, Maryland, recruited to the EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) study (n=385).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total (n=385)</th>
<th>Intervention (n=224)</th>
<th>Control (n=161)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal background</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>37.0 (9.3)</td>
<td>37.6 (9.4)</td>
<td>36.3 (9.0)</td>
<td>.17</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White, n (%)</td>
<td>218 (56.6)</td>
<td>139 (62.1)</td>
<td>79 (49.1)</td>
<td>.04</td>
</tr>
<tr>
<td>Black, n (%)</td>
<td>139 (36.1)</td>
<td>72 (32.1)</td>
<td>67 (41.6)</td>
<td></td>
</tr>
<tr>
<td>Other race, n (%)</td>
<td>28 (7.3)</td>
<td>13 (5.8)</td>
<td>15 (9.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td>.43</td>
</tr>
<tr>
<td>Less than high school graduate, n (%)</td>
<td>177 (46.0)</td>
<td>97 (43.3)</td>
<td>80 (49.7)</td>
<td></td>
</tr>
<tr>
<td>High school graduate or General Educational Development, n (%)</td>
<td>96 (24.9)</td>
<td>60 (26.8)</td>
<td>36 (22.4)</td>
<td></td>
</tr>
<tr>
<td>Some college or greater, n (%)</td>
<td>112 (29.1)</td>
<td>67 (29.9)</td>
<td>45 (28.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Sexual orientation</strong></td>
<td></td>
<td></td>
<td></td>
<td>.36</td>
</tr>
<tr>
<td>Heterosexual or straight, n (%)</td>
<td>260 (67.7)</td>
<td>156 (70.0)</td>
<td>104 (64.6)</td>
<td></td>
</tr>
<tr>
<td>Lesbian, queer, or same gender loving, n (%)</td>
<td>24 (6.3)</td>
<td>11 (4.9)</td>
<td>13 (8.1)</td>
<td></td>
</tr>
<tr>
<td>Bisexual, n (%)</td>
<td>100 (26.0)</td>
<td>56 (25.1)</td>
<td>44 (27.3)</td>
<td></td>
</tr>
<tr>
<td>Homeless in the past 6 months, n (%)</td>
<td>257 (66.8)</td>
<td>151 (67.4)</td>
<td>106 (65.8)</td>
<td>.75</td>
</tr>
<tr>
<td><strong>Sex work history</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time in sex work (years), mean (SD)</td>
<td>13.2 (9.5)</td>
<td>13.3 (10.0)</td>
<td>13.1 (8.9)</td>
<td>.80</td>
</tr>
<tr>
<td><strong>Found clients in the past 6 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Street</td>
<td></td>
<td></td>
<td></td>
<td>.97</td>
</tr>
<tr>
<td>Street-based only, n (%)</td>
<td>124 (32.2)</td>
<td>72 (32.1)</td>
<td>52 (32.3)</td>
<td></td>
</tr>
<tr>
<td>Street or other methods, n (%)</td>
<td>260 (67.7)</td>
<td>152 (67.9)</td>
<td>108 (67.5)</td>
<td></td>
</tr>
<tr>
<td>Web-based or mobile app b, n (%)</td>
<td>95 (36.5)</td>
<td>44 (29.0)</td>
<td>51 (47.2)</td>
<td>.001</td>
</tr>
<tr>
<td>Bar, club, or massage parlor b, n (%)</td>
<td>161 (61.9)</td>
<td>88 (57.9)</td>
<td>73 (67.6)</td>
<td>.17</td>
</tr>
<tr>
<td>Referrals from others (sex workers, pimps or managers, or intimate partners) b, n (%)</td>
<td>152 (58.2)</td>
<td>90 (59.2)</td>
<td>62 (56.9)</td>
<td>.71</td>
</tr>
<tr>
<td><strong>Prevalence of HIV and sexually transmitted infections and risk behaviors, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injected any drug in the past 6 months</td>
<td>223 (57.9)</td>
<td>129 (57.6)</td>
<td>94 (58.4)</td>
<td>.88</td>
</tr>
<tr>
<td>Used heroin in the past 6 months</td>
<td>309 (80.3)</td>
<td>178 (79.5)</td>
<td>131 (81.4)</td>
<td>.64</td>
</tr>
<tr>
<td>Used powdered or crack cocaine in the past 6 months</td>
<td>334 (86.8)</td>
<td>195 (87.1)</td>
<td>139 (86.3)</td>
<td>.84</td>
</tr>
<tr>
<td><strong>Reused injection equipment in the past 6 months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syringes or needles c</td>
<td>68 (28.5)</td>
<td>42 (29.6)</td>
<td>26 (26.8)</td>
<td>.64</td>
</tr>
<tr>
<td>Cookers c</td>
<td>106 (44.5)</td>
<td>62 (43.7)</td>
<td>44 (45.4)</td>
<td>.80</td>
</tr>
<tr>
<td>Cotton c</td>
<td>77 (32.2)</td>
<td>44 (31.0)</td>
<td>33 (34.0)</td>
<td>.62</td>
</tr>
<tr>
<td>Condomless sex with clients in the past week</td>
<td>174 (45.2)</td>
<td>98 (43.8)</td>
<td>76 (47.2)</td>
<td>.50</td>
</tr>
<tr>
<td>Positive HIV rapid test a</td>
<td>20 (5.2)</td>
<td>12 (5.4)</td>
<td>8 (5.0)</td>
<td>.85</td>
</tr>
<tr>
<td>New HIV-positive diagnoses a</td>
<td>3 (0.8)</td>
<td>1 (0.5)</td>
<td>2 (1.2)</td>
<td>N/A^d</td>
</tr>
<tr>
<td>Positive gonorrhea e</td>
<td>59 (15.7)</td>
<td>30 (13.7)</td>
<td>29 (18.6)</td>
<td>.20</td>
</tr>
<tr>
<td>Positive chlamydia e</td>
<td>68 (18.1)</td>
<td>33 (15.1)</td>
<td>35 (22.2)</td>
<td>.08</td>
</tr>
</tbody>
</table>

^a<1% of data missing.
Currently, all participants have moved through their 6-month, 12-month, and final 18-month visit windows. The mobile data collection strategy for EMERALD resulted in the enrollment of 385 FSWs into the cohort portion of the study. We have begun the final cross-sectional surveys, with 4% completed. A total of 11 in-depth interviews have been completed with staff and 16 with clients.

**Discussion**

**Principal Findings**

In the United States, structural interventions aimed at reducing HIV and STIs among FSWs are scarce; to our knowledge, this is the first intervention of its kind in the United States. Innovative and adaptable approaches for linking FSWs to health and social services are required. The findings from the EMERALD evaluation will prove useful for tailoring HIV prevention for FSWs and creating a sustainable community-based intervention. Planned dissemination of study findings includes manuscript publications, conference presentations, and community dissemination (eg, infosheets and community meetings).

The SPARC center has had great success building relationships with FSWs and engaging them in clinical and social services. Much of this success is attributed to the synchronization of street- and van-based outreach efforts that focus on harm reduction education, supply distribution, and referrals with a drop-in center space that has on-site health, social, and legal services tailored to the needs of FSWs and women who use drugs. SPARC has the capacity to provide these wide-ranging services through partnerships with local organizations interested in connecting to the population SPARC serves. SPARC provides space in the drop-in center, and partner organizations use the space to offer clinic hours to potential clients who they might not otherwise be able to access. SPARC clients have greater access to services, and the organizations are able to reach a wider segment of the population they seek to serve.

**Limitations**

Despite successful implementation and the novelty of this intervention in a US setting, we encountered several challenges throughout this study. Although TFSWs experience disproportionately high rates of victimization and are underrepresented in research, a lack of comparable recruitment locations in the intervention and control areas prohibited the inclusion of TFSWs in our study [63,64]. In addition, the onset of the COVID-19 pandemic in March 2020 required adaptations to service provision and data collection. In-person services at the SPARC center were stopped on March 17, 2020. However, because the framework was already in place, the SPARC team was able to quickly pivot to an all-outreach service model when the COVID-19 pandemic forced the closure of the drop-in center.

Regarding the EMERALD evaluation, data collection for the cross-sectional survey was paused at the onset of the COVID-19 pandemic, and data collection was shifted to a telephone format for the remaining 18-month surveys. As a result, we were unable to conduct HIV and STI testing for 27 participants. We plan to reopen in-person services at the SPARC center, resume the 96 in-person cross-sectional surveys, and complete the 19 remaining in-depth interviews when data collection with human subjects is deemed safe and appropriate.

**Acknowledgments**

This work was supported by the National Institute on Drug Abuse under grant R01DA041243; National Institute of Mental Health under grant F31MH118817; and Johns Hopkins University Center for AIDS Research, a National Institutes of Health–funded program, under grant P30AI094189. The authors would like to thank the EMERALD participants, SPARC clients, research study and SPARC staff, the EMERALD study community advisory board, the Baltimore City Health Department, the Behavioral Health Leadership Initiative, Legal Aid of Maryland, and the Maryland Volunteer Lawyers Network. Without their contributions, this project would not have been possible.

**Conflicts of Interest**

SS is an expert witness for the plaintiffs in opioid litigation. The other authors have no competing interests to declare.

**References**


Abbreviations

A-CASI: audio-enhanced computer-assisted self-interviewing
EMERALD: Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity
FSW: female sex worker
JHBSHP IRB: Johns Hopkins Bloomberg School of Public Health Institutional Review Board
LTFU: lost-to-follow-up
RV: recreational vehicle
SAPPHIRE: Sex Workers and Police Promoting Health in Risky Environments
SPARC: Sex Workers Promoting Action, Risk Reduction, and Community Mobilization

STI: sexually transmitted infection

TFSW: transgender female sex worker

Edited by G Eysenbach; submitted 23.10.20; peer-reviewed by P Nguyen, J Xu, J Mitchell, R Poluru; comments to author 08.12.20; revised version received 29.01.21; accepted 16.02.21; published 16.04.21.

Please cite as:
Silberzahn BE, Tomko CA, Clouse E, Haney K, Allen ST, Galai N, Footer KHA, Sherman SG
The EMERALD (Enabling Mobilization, Empowerment, Risk Reduction, and Lasting Dignity) Study: Protocol for the Design, Implementation, and Evaluation of a Community-Based Combination HIV Prevention Intervention for Female Sex Workers in Baltimore, Maryland
JMIR Res Protoc 2021;10(4):e23412
URL: https://www.researchprotocols.org/2021/4/e23412
doi:10.2196/23412
PMID:33861210

©Bradley E Silberzahn, Catherine A Tomko, Emily Clouse, Katherine Haney, Sean T Allen, Noya Galai, Katherine H A Footer, Susan G Sherman. Originally published in JMIR Research Protocols (http://www.researchprotocols.org), 16.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on http://www.researchprotocols.org, as well as this copyright and license information must be included.
Nutritional Implications of Baby-Led Weaning and Baby Food Pouches as Novel Methods of Infant Feeding: Protocol for an Observational Study

Rachael W Taylor¹, BSc (Hons), PhD; Cathryn A Conlon², BSc (Hons), MMEdSci, PhD, RGN; Kathryn L Beck², BSc, BPhEd, MSc, PGDipDiet, PhD; Pamela R von Hurst², BSc (Hons), PhD; Lisa A Te Morenga³, BSc, PhD; Lisa Daniels¹, BSc, MDiet, PhD; Jill J Haszard⁴, BSc, BBiomedSci, MSc, PGDip (Biostats), PhD; Alison M Meldrum⁵, BDS, MDS; Neve H McLean⁶, BSc, MDiet; Alice M Cox¹, BSc, MDiet; Lesieli Tukuafu⁵, BDS; Maria Casale², BSc, MSc; Kimberley J Brown², BSc, MSc; Emily A Jones⁵, MSphLangTher, BAAppSci (SpPath); Ioanna Katiforisi⁶, BA, BSc; Madeleine Rowan⁶, BSc; Jenny McArthur¹, BEd; Elizabeth A Fleming⁶, MCApSc; Ben J Wheeler⁷, MBChB, PhD; Lisa A Houghton⁶, PhD; Aly Diana⁸, MD, PhD, MPH; Anne-Louise M Heath⁶, BA (Hons), BSc (Hons), PhD

¹Department of Medicine, University of Otago, Dunedin, New Zealand
²School of Sport, Exercise and Nutrition, College of Health, Massey University, Auckland, New Zealand
³Research Centre for Hauora and Health, Massey University, Wellington, New Zealand
⁴Biostatistics Centre, University of Otago, Dunedin, New Zealand
⁵Faculty of Dentistry, University of Otago, Dunedin, New Zealand
⁶Department of Human Nutrition, University of Otago, Dunedin, New Zealand
⁷Department of Women's and Children's Health, University of Otago, Dunedin, New Zealand
⁸Nutrition Working Group, Faculty of Medicine, Universitas Padjadjaran, Bandung, Indonesia

Corresponding Author:
Anne-Louise M Heath, BA (Hons), BSc (Hons), PhD
Department of Human Nutrition
University of Otago
70 Union St West
Dunedin, 9054
New Zealand
Phone: 64 34798379
Email: anne-louise.heath@otago.ac.nz

Abstract

Background: The complementary feeding period is a time of unparalleled dietary change for every human, during which the diet changes from one that is 100% milk to one that resembles the usual diet of the wider family in less than a year. Despite this major dietary shift, we know relatively little about food and nutrient intake in infants worldwide and virtually nothing about the impact of baby food “pouches” and “baby-led weaning” (BLW), which are infant feeding approaches that are becoming increasingly popular. Pouches are squeezable containers with a plastic spout that have great appeal for parents, as evidenced by their extraordinary market share worldwide. BLW is an alternative approach to introducing solids that promotes infant self-feeding of whole foods rather than being fed purées, and is popular and widely advocated on social media. The nutritional and health impacts of these novel methods of infant feeding have not yet been determined.

Objective: The aim of the First Foods New Zealand study is to determine the iron status, growth, food and nutrient intakes, breast milk intake, eating and feeding behaviors, dental health, oral motor skills, and choking risk of New Zealand infants in general and those who are using pouches or BLW compared with those who are not.

Methods: Dietary intake (two 24-hour recalls supplemented with food photographs), iron status (hemoglobin, plasma ferritin, and soluble transferrin receptor), weight status (BMI), food pouch use and extent of BLW (questionnaire), breast milk intake (deuterium oxide “dose-to-mother” technique), eating and feeding behaviors (questionnaires and video recording of an evening meal), dental health (photographs of upper and lower teeth for counting of caries and developmental defects of enamel), oral motor skills (questionnaires), and choking risk (questionnaire) will be assessed in 625 infants aged 7.0 to 9.9 months. Propensity
score matching will be used to address bias caused by differences in demographics between groups so that the results more closely represent a potential causal effect.

**Results:** This observational study has full ethical approval from the Health and Disability Ethics Committees New Zealand (19/STH/151) and was funded in May 2019 by the Health Research Council (HRC) of New Zealand (grant 19/172). Data collection commenced in July 2020, and the first results are expected to be submitted for publication in 2022.

**Conclusions:** This large study will provide much needed data on the implications for nutritional intake and health with the use of baby food pouches and BLW in infancy.

**Trial Registration:** Australian New Zealand Clinical Trials Registry ACTRN12620000459921; http://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=379436.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/29048


**KEYWORDS**
infant; diet; complementary feeding; food pouch; baby-led weaning; iron; growth; eating behavior; feeding behavior; dental health; choking; breast milk

**Introduction**

**Background**
The biggest dietary change in every human’s life is the change from consuming a 100% milk diet in the first months of life to consuming a diet that is broadly the same as that of the rest of the family by the first birthday. This change is large. In fact, an infant following the current infant feeding guidelines globally [1-4] would have breast milk (or infant formula) as the sole source of nutrition until around 4 to 6 months of age. The first “complementary” foods would then be introduced one at a time in 1 to 2 teaspoon serves of “thin smooth purée,” followed by a progression in food texture from puréed to mashed to chopped, and by around 1 year of age, family foods are consumed. While many countries have reasonably up-to-date information on the nutrient intake of infants [5-7], there is no such comparable data in New Zealand infants. Additionally, no country has specific information yet on the impact of the revolution in infant feeding offered by the new phenomena of baby food “pouches” and “baby-led weaning” (BLW).

Baby food pouches are squeezeable containers with a plastic spout described as a “mess-free and easy alternative for baby food on the go” [8]. They have immense appeal to parents for their convenience [9] and perceived superior safety and freshness over more traditional glass jars [9], and parental perceptions that they are healthy and enjoyed by the baby. This appeal is demonstrated by an extraordinary market share. A recent analysis of 24 major brands of infant and toddler foods in the United States (that represent more than 95% of market share) showed that 56% of the food products were packaged as squeeze pouches [10]. The market share continues to grow. Sales of baby food pouches in Europe have grown annually by 125% in Spain and 916% in the Ukraine [11]. Surprisingly, there appears to have been almost no direct research on the possible impact of this new technology on infant diet or health, despite several groups cautioning against the use, highlighting the urgent need for research on the safety, nutrition, and health impacts of pouches [12-14].

Although the foods offered in baby food pouches are broadly similar in content to those offered as baby foods in jars and cans, there has been some concern, albeit not universal [15], that the content of added sugar might be higher [10,13,14,16]. Certainly, cereal products in pouches are a poor source of iron and should not be used to replace iron-fortified infant cereals [15]. Regardless of nutrient content, the delivery method itself has the potential to markedly change infant nutrition for several reasons. Anecdotal reports suggest food products are being consumed straight from the pouch, unsupervised, and “on the go,” so they are outside usual eating contexts. This could have a number of important impacts on infant health and development.

First, these smooth highly processed products with multiple blended ingredients bear little resemblance to intact fruits and vegetables, and they are marketed well beyond the early weeks of complementary feeding when it might be argued a “super smooth” product is appropriate (eg, many are marketed for infants 8 months plus). This raises a number of questions. Do these products increase energy intake because they are so easy to eat (smooth consistency and do not require chewing; an entire 120-g serve can be accessed merely by squeezing the soft pouch and sucking)? Conversely, does the ease of consumption lead to displacement of other more nutrient-rich foods, such as breast milk (or infant formula), from the diet (eating “on the go”) is unlikely to be consistent with New Zealand Ministry of Health recommendations that until 8 months of age infants are offered solid foods after milk to avoid displacing nutrient-rich milk [2]? Certainly, the structure of baby food pouches is described as “facilitating rapid passage of solid foods” [12], which suggests overeating is possible.

Second, if these pouches are indeed being consumed “on the go” by infants, what implications does this have for learning about food and eating? For example, do infants who feed themselves by sucking a purée out of a pouch while “on the go” have the same relationship with food as that for infants who sit and eat with their family? The description of pouches “promoting self-feeding and independence” [11] may suggest not. Positive reciprocal interaction during feeding or “responsive feeding” (ie, the caregiver and child respond appropriately to one another’s cues) may support the infant’s innate ability to respond to hunger and satiety cues [17]. If pouch use promotes...
“self-feeding and independence,” does frequent pouch use result in a lack of interaction during feeding? Moreover, is there any association between regular pouch use and the infant’s satiety responsiveness (ie, ability to cease eating once full)? Furthermore, parents are able to model healthy eating behaviors when the infant takes part in shared eating occasions, and this is associated with infants being less “food responsive” (ie, less likely to eat just because food is available) [18]. Beyond this, family routines and rituals around shared meals provide “a predictable structure that guides behavior and an emotional climate that supports early development” [19]. If feeding interactions and modeling influence the development of eating behaviors, then regularly eating from a pouch while “on the go” rather than in a shared eating setting may have important impacts on the development of infant eating behaviors.

Third, what is the impact of prolonged exposure to these often sweet and acidic (and therefore presumably cariogenic) foods on erupting teeth? There is a strong relationship between the frequency of cariogenic food intake and childhood caries [20]. Children who experience caries as infants or toddlers (ie, early childhood caries) have a much greater risk of developing caries in their permanent teeth [21], with children who have high exposure to sugars during infancy having a much greater risk of dental caries at 3 years than children who have less exposure to sugars in infancy [22]. Frequent consumption of sweet and acidic foods in early infancy may be of particular concern because newly erupted teeth have immature enamel and are more likely to develop caries [21]. Dental caries can have immediate and ongoing impacts on child health and quality of life, including reduced weight gain if food consumption is impacted, pain and discomfort, altered sleeping habits, and in the worst cases, hospitalization. Concerns have been raised about the possible impact of baby food pouches on pediatric dental health [12,23]. Certainly, advice to limit cariogenic foods to meal times [24] is not being followed if fruit- or cereal-containing pouches are used for snacks while “on the go” and therefore between meals.

While the use of food pouches is starting to be investigated internationally [25], there is no published research examining their relationship with infant nutrient intake, eating and feeding behaviors, growth, or dental health. The US Feeding Infants and Toddlers Study (FITS 2016) has collected but not yet reported data on the use of baby food pouches [25], but FITS 2016 did not collect data on nutritional status or health outcomes. Despite the lack of research on the possible impacts of pouch use, health professionals have expressed concerns [23,26-29]. A recent New York Times article [26] reports a spokeswoman for the American Academy of Pediatrics expressing concerns that pouch use may lead to children “overiding their body’s own cues for hunger and fullness” and recommending families should have established times for meals rather than “pouching the calories throughout the day.” In response to the article, a pediatric occupational therapist describes behaviors she has observed as follows: “Pouches appear to solve so many problems: kids who make a mess, kids who refuse fruits and veggies, kids who refuse to touch food or use a utensil, kids who won’t sit still through a meal… I see toddlers waltzing through homes every day, sucking on pouches to start, end, or replace a meal” [26].

These are just anecdotal reports, but they underline the urgent need to determine the effect of pouches on infant nutrition and health. Interestingly, given the lack of research in this area, some health professionals in Germany [27] and the United Kingdom [23] have already gone so far as to recommend against the use of baby food pouches.

The second recent phenomenon in infant feeding is the popular adoption of BLW, an alternative approach to introducing solids to infants. In BLW, infants feed themselves all their foods from the start of the “complementary feeding” period. This means no spoon feeding by a parent, and only “finger foods” are offered [30]. BLW differs considerably from the more traditional approach espoused by infant feeding guidelines in many countries [1-4], in which the infant gradually learns how to eat solid foods safely by eating foods with progressively increasing textures from pureéd to mashed to chopped to whole. We currently do not know how pervasive BLW is, although a recent New Zealand study [31] suggests more than half of families have tried it, with approximately 30% following it regularly. However, considerable concern has been expressed by health professionals about the potential increased risks of iron deficiency, growth faltering, and choking with BLW [32,33].

The limited international research base would suggest these concerns may be justified. We have shown previously in a small sample of infants aged 6 to 8 months that those following BLW had a much lower iron intake than traditionally fed infants [34], which is an issue given that iron is already a nutrient of concern in infants and toddlers, both in New Zealand [35] and internationally [36]. In order to truly know whether concern about low iron intake in infants following BLW is justified, the biochemical iron status of infants must be determined as iron intake is a notoriously poor indicator of iron status [37]. Only two small studies appear to have examined intake of other nutrients in infants following BLW, suggesting that intakes of zinc and vitamin B12 may be lower [34], and intakes of total fat, saturated fat [34], and sodium [38] may be higher than those for traditional spoon feeders. These important differences require clarification in a much larger sample. Proponents of BLW argue that infants are able to feed themselves sufficient food from 6 months of age and that allowing children to have control over their own eating promotes a greater ability to regulate their own appetite appropriately. Whether this is indeed true or translates into differences in growth rates is uncertain given that the few existing studies [39,40] have used parental reports, which can be inaccurate [41], particularly in infants who are growing so rapidly. Lastly, foods, such as raw apple, raw carrot, and grapes are some of the most commonly seen foods in videos promoting BLW, despite the substantial choking risk they pose to young children [42]. Whether choking rates differ in BLW versus more traditional solid feeding is not clear. Only one large study internationally has specifically investigated choking rates in infants following BLW rather than traditional spoon feeding (TSF) [43]. This study suggested that choking rates may in fact be lower for infants who consume finger foods regularly, but it recruited the BLW participants from BLW websites rather than the general population. This may have biased the results.
as choking and gagging are common topics discussed on BLW websites (particularly the importance of not mistaking gagging for choking), and this may have influenced the reporting of choking rates in the BLW infants. Given the substantial number of parents following this approach with their infants and BLW’s widespread online presence (7,690,000 results on Google; December 24, 2020), it is critically important to determine the health risks of BLW so that health professionals and policy makers can provide families with evidence-based advice on how to feed their infants safely.

Infant milk (breastmilk or infant formula) is a substantial component of the diet for infants during the complementary feeding period, providing more than half of their energy intake at 7 months of age [44]. While it is straightforward to estimate intake of infant formula for those who consume it, researchers currently have to use a “one size fits all” approach to estimate breast milk intake, either using a single volume for all breastfed infants of a particular age [45] or excluding breastfed infants from dietary analyses [46]. Neither approach is ideal because breast milk intake varies considerably between mother-infant pairs [47] and because breastfed infants do not necessarily have the same food intake or socioeconomic background as formula fed infants. The conventional method used to measure breast milk intake is to weigh the infant before and after every feed. However, this “test-weighing” technique is time consuming and may disturb usual feeding patterns. In contrast, the stable isotope deuterium oxide technique requires the mother to consume a small amount of the stable isotope (deuterium oxide) in water, and the amount of this marker transferred to the infant (ie, via the breast milk the infant consumes) is then measured by collecting saliva samples from the mother and infant over the following fortnight [48-50]. The normal feeding pattern is not disturbed, and the total volume of breast milk consumed by the infant over the fortnight can be accurately assessed. The use of this technique will allow the First Foods New Zealand (FFNZ) study to collect data that will enable more accurate estimates of nutrient intake in this age group and allow us to generate predictive models that use infant and diet characteristics that are routinely measured to estimate breast milk volumes. Such models would be invaluable for estimating total nutrient intake for breastfeeding infants (69% of infants at 4-8 months in New Zealand [51]) in future studies.

The FFNZ study will determine the iron status, growth, food and nutrient intakes, breast milk intake, eating and feeding behaviors, dental health, oral motor skills, and choking risk of New Zealand infants, with a particular focus on the use of baby food pouches and BLW.

**Primary Objective**

In infants aged 7.0 to 9.9 months, we will determine whether iron status and BMI z-score differ according to the extent of food pouch use and complementary feeding approach (BLW compared with TSF).

**Secondary Objectives**

In infants aged 7.0 to 9.9 months, we will estimate the following: (1) Nutrient intake, nutrient adequacy, and foods of cultural importance in New Zealand infants and in infants fed using baby food pouches regularly or those following BLW; (2) Breast milk intake in New Zealand infants and in infants fed using baby food pouches regularly or those following BLW; (3) Prevalence and nature of food pouch use; (4) Prevalence of BLW; (5) How eating behaviors (ability to eat with appetite, speed of eating, and picky eating) and feeding behaviors (parental responsiveness to infant hunger and satiety cues) differ according to the extent of food pouch use and complementary feeding approach (BLW compared with TSF); (6) How dental health differs according to the extent of food pouch use and complementary feeding approach (BLW compared with TSF); (7) How oral motor skills differ according to the extent of food pouch use and complementary feeding approach (BLW compared with TSF); and (8) How the risk of choking differs according to the extent of food pouch use and complementary feeding approach (BLW compared with TSF).

**Methods**

**Design**

The FFNZ study is an observational cross-sectional study of food and health in infants aged 7.0 to 9.9 months. The study will compare infants using baby food pouches with those not using these pouches, and those following BLW with those following TSF, while collecting data on nutrient intake and nutritional status in this age group in general. The age range has been chosen because it is close enough to when complementary feeding starts (usually 4-6 months of age) that we can expect to see large variations in both baby food pouch use and BLW rates, while also giving enough time from the start of complementary feeding for eating patterns to have had an impact on iron status and growth. A narrow age range has specifically been chosen because diet changes rapidly in infancy. Observational study designs are appropriate for identifying associations between behaviors as they are carried out in the “real world.” While a randomized controlled trial is required to determine causality, it is not ethical to randomize infants to follow BLW or pouch use because that would require randomization of participants to eating patterns that health professionals have concerns about [23,26-28,32,33]. Instead, we will use propensity score matching [52], which is able to remove some of the bias caused by differences in demographics between groups so that the estimates of the impact of pouch use or BLW on infant diet and health will more closely represent a potential causal effect [53].

**Participants and Recruitment**

In total, 625 parents/guardians who have an infant less than 9.9 months of age will be recruited from two regions of New Zealand (Dunedin and Auckland) to participate in the study when their infant is aged 7.0 to 9.9 months. Recruitment will occur by advertisement and word of mouth and will target all infants rather than those adopting BLW, TSF, or food pouch use. We aim to recruit a sample that is broadly representative of the ethnicity and socioeconomic status of New Zealand children. It is not feasible to recruit a truly representative sample using typical methods, such as electoral roll and door knocking, because they would identify very few infants in the narrow age band that is necessary for this study (because diet changes so
rapidly in infancy). We will, however, collect data from a diverse range of ethnic and socioeconomic groups by (1) engaging with Māori and Pasifika community health organizations to assist with recruitment, (2) targeting recruitment in suburbs with a high proportion of Māori, Pasifika, and Asian populations, and (3) having research team members who have experience working with or who culturally identify with Māori, Pasifika, and Asian communities. We will also statistically weight the estimates to account for demographic disparities if appropriate. The study has ethical approval from the Health and Disability Ethics Committees New Zealand (19/STH/151), and written informed consent will be obtained at the first appointment. The study is registered with the Australian New Zealand Clinical Trials Registry (registration number: ACTRN12620000459921).

Sample Size

Our sample size calculation is based on comparing the BMI z-score and plasma ferritin concentration in infants following BLW and TSF, as there are currently no data internationally on these measures in pouch users. Our recent studies suggest that 29% of infants will meet the definition of BLW [31] and 70% of enrolled infants will provide a blood sample [54]. A recruitment size of 625 would therefore enable us to collect complete data from 125 BLW and 312 TSF infants, which would be sufficient to detect a difference of 0.3 for the BMI z-score and a 5-μg/L lower plasma ferritin concentration in the BLW group assuming a mean of 29 μg/L in the TSF group [54], both with 80% power and α of .05. As outlined above, we expect pouch use to be very common, based on the 70% market share they have, but we do not have the data needed to calculate the sample size required to detect differences in health outcomes with pouch use. With a sample size of 625, however, we will be able to estimate the prevalence of frequent pouch use to a 95% precision level of at least ±4%.

Data Collection

Overview

Participation in the study will involve three (participants in the main study) or five (participants in a consecutive subsample of breastfed infants) contacts over 2 weeks following recruitment. For the main study (n=625), the first main appointment will generally be held in the participant’s home and involve a 24-hour diet recall, completion of two questionnaires, and anthropometric measurements of the child. The second main appointment will generally take place at university research rooms, and involve a second 24-hour diet recall and photography of the infant’s teeth to assess dental health. A third main appointment will take place at our university rooms (Dunedin) or a local blood testing facility (Auckland) to collect a blood sample to measure the iron status. Finally, a self-administered questionnaire will be completed by the participants in their homes. For the subsample (n=150) involved in the measurement of breast milk intake, a stable isotope will be given at the first main appointment, with three additional saliva samples collected over the ensuing fortnight (the third sample being collected at the second main appointment).

Demographic Data

At the initial appointment, ethnicity, maternal education, maternal work status, household deprivation (New Zealand deprivation index 18 [55]), household food security [56], and childcare use will be collected by questionnaires, using the New Zealand census questions where relevant. These data will be used to describe the sample and minimize bias.

Measuring Baby Food Pouch Use

This is related to primary objective 1 and secondary objectives 1, 2, 3, 5, 6, 7, and 8. We will measure the frequency of pouch use in the past month by a questionnaire. We intend to define infants as being frequent baby food pouch users if their parents state that they are currently being given food from a food pouch “5 to 6 times a week,” “once a day,” or “more than once a day,” although this may need to be modified when the distribution of intakes is determined (there are currently no published data on the frequency of baby food pouch use to base this cutoff on). We will collect data on the frequency of pouch use, use of “ready-to-eat” pouches versus “home-filled” pouches, extent to which the infant feeds directly from the pouch rather than being fed by spoon, types of foods given in “pouches,” contribution of pouch foods to total intake of solid foods, proportion of the “pouch” consumed on a typical eating occasion, duration of a typical eating occasion, physical situations in which pouches are used, proximity of an adult when the pouch is being used, reasons for using pouches rather than other methods of food delivery, and anything not liked about using baby food pouches. Key pouch questions will be asked referring to when the infant first started eating solids, when the infant was around 6 months of age, and “now.”

Measuring BLW

This is related to primary objective 1 and secondary objectives 1, 2, 4, 5, 6, 7, and 8. Parents will be asked to describe the way their infants were fed when they first started eating solids, when they were around 6 months of age, and “now,” using five answer options. Parents who choose “spoon fed by an adult” or “mostly spoon fed by an adult, some baby feeding themselves” will be classified as TSF. Parents who select “about half spoon feeding by an adult and half baby feeding themselves” will be classified as partial BLW [31]. Those who report “mostly baby feeding themselves, some adult spoon feeding” or “baby feeding themselves” will be assigned to full BLW [57,58]. As there is no validated definition of BLW, these definitions have been designed to capture the major point of difference between BLW and TSF, while allowing occasional adult spoon feeding.

Iron Status

This is related toprimary objective 1. A nonfasting venipuncture blood sample will be collected at the third main appointment (3-mL EDTA anticoagulated vacutainer blood collection tube; Becton Dickinson and Company) to determine the plasma ferritin concentration and iron status defined using the body iron concentration (calculated using plasma ferritin and soluble transferrin receptor concentrations [36]) and hemoglobin concentration (from a complete blood count). The iron status categories are defined in Table 1 [54].
Table 1. Iron status categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Body iron value</th>
<th>Hemoglobin value</th>
<th>Plasma ferritin value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iron sufficient</td>
<td>≥0 mg/kg</td>
<td>≥105 g/L</td>
<td>≥15 μg/L</td>
</tr>
<tr>
<td>Iron depleted</td>
<td>≥0 mg/kg</td>
<td>≥105 g/L</td>
<td>&lt;15 μg/L</td>
</tr>
<tr>
<td>Early functional iron deficiency</td>
<td>&lt;0 mg/kg</td>
<td>≥105 g/L</td>
<td>N/A(^a)</td>
</tr>
<tr>
<td>Iron deficiency anemia</td>
<td>&lt;0 mg/kg</td>
<td>&lt;105 g/L</td>
<td>N/A</td>
</tr>
</tbody>
</table>

\(^a\)N/A: not applicable.

As ferritin is an acute phase reactant and can be artificially elevated by inflammation, we will also analyze two inflammatory markers (C-reactive protein and α-1-acid glycoprotein) as recommended by the World Health Organization (WHO) [59]. This will enable us to use the BRINDA (Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia) statistical method to adjust for inflammation on a continuous scale, consistent with the assumption that higher concentrations of inflammatory markers will be associated with a greater effect on plasma ferritin concentration [60].

Participants will be given verbal instructions on how to apply a local anesthetic (Ametop gel; Perstorp Pharma) and will be given access to an instruction video. The gel is to be applied to the insides of both of the infant’s elbows (this enables phlebotomy to be attempted on the second arm if necessary) and covered with an occlusive dressing at least 1 hour (no more than 4-6 hours) before the blood test appointment. It is removed after 30 to 45 minutes. The blood sample will be taken by a pediatric phlebotomist. The research team has extensive experience overseeing research projects involving the collection of venipuncture blood samples from infants and toddlers [54,60,61], with 70% [54] to 92% [61] of participants providing samples.

Commercial laboratories (Southern Community Laboratories, Dunedin, New Zealand and Labtests NZ, Auckland, New Zealand) will determine complete blood count (requires fresh blood) and plasma ferritin, so that the iron status can be immediately communicated to the infant’s general practitioner if the infant is identified as having anemia. The remaining plasma will be frozen at −80°C for batch analysis of soluble transferrin receptor, C-reactive protein, and α-1-acid glycoprotein concentrations at the University of Otago Department of Human Nutrition Laboratory at the end of the study [54].

**Anthropometry**

This is related to primary objective 1. Infant weight will be measured at the initial appointment on an electronic scale (model 354; Seca) and length will be measured on a 99-cm measuring mat (model SE210; Seca) in duplicate following WHO protocols [62]. BMI (weight in kg divided by height in meters squared) will be calculated, and BMI-for-age z-scores will be determined using WHO reference data [63]. We will also request consent from parents to collect information on BMI from the B4 School Check [64] when the children are 4 years of age. This will enable us to look at the effects of pouches and BLW on growth longitudinally.

**Infant Diet**

This is related to secondary objective 1. Information on infant nutrient intake and adequacy, food group intake, dietary patterns, and culturally important foods will be obtained using interviewer-administered multiple pass 24-hour recalls collected at the first and second appointments. The two 24-hour recalls take place on different days of the week to capture variation in intake between days. Collecting two 24-hour recalls will enable us to calculate “usual intake” using the multiple source method (MSM) for estimating usual dietary intake of individuals [65]. Photograph prompts will be used to assist recall of foods eaten. Participants will be asked to photograph (using their own smartphone or a camera provided) the eating surface (eg, plate, high chair surface, and table) at the start of all meals and snacks from midnight to midnight on the day before the appointment. The quality of the photographs will not be important as long as they are clear enough to remind the parent what the child ate. Diet recalls will be analyzed with FoodWorks (version 10, Xyris Software) using the New Zealand Food Composition database FOODfiles 2018 Version 01 [66]. Nutrient information for commercial infant foods and milk will be determined by generating recipes using the ingredients lists on food products modified to match the nutrient information panel on the packet so that the contribution of nutrients that do not appear on the nutrient information panel can be included [15]. Information on supplement use will be collected by a questionnaire.

Particular focus will be on free sugars and added sugars given a recent small study [45] suggested that even by 7 months of age, 12% of New Zealand infants may already be consuming free sugars at levels that are above the WHO recommendation [67]. Parents are discouraged from adding sugar to infants’ diets because it is unnecessary and may increase liking of sweet foods [2]. In addition, the WHO recommends that free sugars should be <5% of energy intake due to the dose-response relationship between free sugar intake and dental caries (even in populations with water fluoridation) [67]. Data on free sugars and added sugars are available in the New Zealand food composition database [66,68].

Questionnaire and 24-hour recall data will also be used to determine the extent to which key indicators of diet quality are being met, as guided by the New Zealand Ministry of Health Eating and Activity Guidelines for New Zealand Infants and Toddlers [2].

**Breast Milk Intake**

This is related to secondary objectives 1 and 2. We will obtain accurate data on the amount of breast milk infants consume using the stable isotope (deuterium oxide) “dose-to-mother”...
technique [48-50] in a consecutive sample of 150 breastfeeding mother-infant dyads. This will enable us to measure intake to ±34 mL/day (ie, 5% of expected total intake [47]) at a 95% precision level. The isotope will be administered to the mother orally at the initial appointment (after collection of the baseline saliva sample), and saliva samples will be collected from the mother and infant at three further appointments to measure the disappearance of the deuterium from the mother and appearance in the infant. Baseline and three postdose sampling points (days 2-3, 7-9, and 13-14) are required to achieve adequate accuracy and precision of human milk intake. Height/length and weight will be measured at baseline, with weight measured again at the final appointment for both mother and infant, so that breast milk intake can be calculated. Breast milk intake data will be used along with questionnaire and recall data to generate predictive equations of breast milk intake so that intake can be estimated for participants who did not have breast milk intake measured.

Eating Behaviors

This is related to secondary objective 5. Eating behaviors will be assessed using the following four subscales from the Children’s Eating Behavior Questionnaire (CEBQ) [69]: “satiety responsiveness” (eating appropriately in response to appetite), “food responsiveness,” “food enjoyment” (eating in response to environmental food cues rather than hunger), and “slowness in eating.” Although a Baby Eating Behavior Questionnaire has been developed [70], it is designed for infants who are exclusively milk fed, so it would not capture complementary feeding. Food fussiness will be measured using five items in the “picky eating” subscale of the Toddler-Parent Mealtime Behavior Questionnaire [71]. We have demonstrated the internal consistency and reliability of these scales in New Zealand infants at 12 months of age, with Cronbach α ranging from .83 to .90 [58]. Participants will also be asked whether they feed their infants any foods of particular cultural relevance. Feeding behaviors will be determined by observing how infants eat and how parents react in response to hunger and satiety cues by videotaping one evening meal (at which solid foods are offered) in each infant recruited in the Dunedin cohort (n is approximately 300). Participants will be issued a GoPro wide-angle video camera (Hero 2018; GoPro Inc) and tripod at their first appointment and asked to video record the main meal on the day for which the second 24-hour recall will be obtained. The camera will take a continuous video from approximately 10 minutes before the infant first joins the meal to when they leave it. Videos will be coded using the Responsiveness to Child Feeding Cues Scale [72].

Dental Health

This is related to secondary objective 6. Photographs of the infant’s upper and lower teeth will be taken by trained interviewers using a dedicated study Oppo Reno2 Z (Oppo) mobile phone with a small portable Smile Light MDP lighting source specifically designed for taking dental pictures [73]. These images will be examined by a single registered dental practitioner, with blinded evaluation of a subset by another examiner, using validated indices for caries and developmental defects of enamel, which have a positive correlation with dental caries [74].

Oral Motor Skill Development

This is related to secondary objective 7. The questionnaires administered at the final appointment include the validated Child Oral Motor Proficiency Scale (ChOMPS) to identify oral motor and eating skill delay [75,76] and the Pediatric Eating Assessment Tool (PediEAT) to measure behaviors that characterize symptoms of feeding difficulties [77,78]. Both questionnaires have age-based reference values for infants and rely on parent reporting [79,80].

Choking

This is related to secondary objective 8. The questionnaire administered at the initial appointment will include questions on choking since birth. We developed these retrospective questions for previous work in this age group and have demonstrated that they provide data that are comparable to choking data collected prospectively using a daily choking calendar [81].

Statistical Analysis

We expect that there will be crossover between BLW/TSF status and pouch use. We will be able to explore whether mean differences in energy and nutrient intake, as well as other measures, between pouch users and nonusers are different for those who use BLW and those who do not. These results will be stratified by BLW/TSF status, and estimated differences will be compared. Regression models will be used to determine differences between groups. Propensity score matching will be undertaken to reduce the bias caused by differences in demographics between the groups (eg, maternal education and ethnicity), infant age (to the nearest week), and sex. Propensity score matching is not like traditional paired matching, where each individual is matched to another individual in the other group according to covariates. Instead, propensity score matching uses a participant’s propensity score (found using covariates) and estimates what their outcome (eg, energy intake) would have been if they were in the other of two dichotomous groups. By using this method, we expect the estimates will more closely represent a potential causal effect [53]. Another advantage is that it allows data from the whole sample to be used, unlike a traditional matched analysis in which only matched pairs are analyzed.

Estimates of BLW, frequent pouch use, nutrient intake, status, and adequacy will be calculated for the whole sample along with 95% CIs. If the sample is not demographically representative of the wider population, statistical weighting of these estimates will be undertaken using the survey command in Stata (StataCorp).

Nutrient intake will be determined using 24-hour recall data adjusted to provide estimates of usual intake (using the MSM method [65]). Adequacy of intake will be determined as follows: for zinc (for which an estimated average requirement [EAR] is available for New Zealand and Australia [82]), the EAR cut point method will be used; for iron (for which an EAR and
tables of probabilities of inadequate intakes are available for the United States of America and Canada [83]), the full probability approach will be used [83]; for other nutrients (only an adequate intake [AI] is available for this age group in New Zealand and Australia [82]), mean group intake above the AI will be considered to indicate adequacy, but a conclusion as to inadequacy will not be possible if mean group intake is below the AI [84]. The BRINDA method [60] will be used to adjust plasma ferritin, and therefore, body iron and iron status, for the impact of inflammation.

The best-fitting polynomials to predict breast milk intake will be estimated by regression models using fractional polynomial functions of variables, such as age, sex, body weight, and food and beverage intake (eg, kJ/day). This will result in equations that can be used to predict breast milk intake based on a variety of input variables. Ideally one based on data that can be collected in a single clinical appointment, and another that uses data requiring more extensive collection in a research or surveillance setting.

**Results**

This observational study has full ethical approval from the Health and Disability Ethics Committees New Zealand (19/STH/151) and was funded in May 2019 by the Health Research Council (HRC) of New Zealand (grant 19/172). Data collection commenced in July 2020, and the first results are expected to be submitted for publication in 2022. Data collection will only take place while New Zealand is in Alert Levels 1 or 2 during the COVID-19 pandemic. The Otago and Auckland regions of New Zealand, where the data collection will take place, have been in Level 1 for all but 7 weeks in Otago (all at Level 2) and 11 weeks in Auckland (7 weeks at Level 2 and 4 weeks at Level 3) since July 2020, as overall case numbers in New Zealand remain extremely low (<2000 in a population of more than 5 million). As daily life is essentially normal in Level 1 with the exception of closed international borders and Level 2 just requires some physical distancing and appropriate hygiene recommendations, we feel confident that the pandemic will have relatively little effect on our data.

**Discussion**

This large observational study will provide much needed data on nutrient intake (including breast milk intake) and nutritional status (specifically iron status, growth, and dental health) in a large diverse sample of New Zealand infants. However, our data will also have considerable international appeal given the lack of research assessing the implications for nutritional intake and health for those infants who obtain a large proportion of their food via baby food pouches. Similarly, determining how iron status, growth, nutrient intake, and choking risk may differ in infants following BLW compared with TSF is urgently warranted given the widespread interest in this alternative approach to complementary feeding worldwide.

**Acknowledgments**

We acknowledge the wider members of the research team (Owen Mugridge, Rio Monzales, Hayley Dodd, and Glenna Paterson), the support of Southern Community Laboratories, and the participants. This study is supported by the Health Research Council (HRC) of New Zealand (19/172). RWT is supported by a Fellowship from the Karitane Products Society. NHM and KJB are supported by PhD scholarships from the Health Research Council of New Zealand, AC is supported by a PhD scholarship from the University of Otago, and MC is supported by a PhD scholarship from Massey University. The HRC had no role in study design, writing of the protocol manuscript, or the decision to submit the manuscript for publication.

**Authors’ Contributions**

RWT and ALMH are the co-principal investigators of the First Foods New Zealand study. RWT, CAC, KLB, PRvH, LAT, LD, JJH, AMM, LAH, and ALMH designed the project and applied for funding. RWT and ALMH produced the first and subsequent drafts of the manuscript. JJH advised on study design, sample size calculation, and statistical analysis. JA is the project coordinator. LD, NHM, AC, LT, MC, KB, EJ, IK, EF, and DA developed the study data collection protocols, and NHM, AC, LT, MC, KB, EJ, IK, MR, and JM undertook data collection. All authors made important intellectual contributions to the manuscript, and all have read and approved the final version.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**
Peer-reviewer reports and grant approval from Health Research Council of New Zealand.

[PDF File (Adobe PDF File), 1958 KB - resprot_v10i4e29048_app1.pdf ]

**Multimedia Appendix 2**
Ethics approval.

[PDF File (Adobe PDF File), 396 KB - resprot_v10i4e29048_app2.pdf ]

https://www.researchprotocols.org/2021/4/e29048
References


73. Smile Lite MDP. Smile Line. URL: https://www.smileline.ch/en/products/smile-lite-mdp [accessed 2021-04-06]


Abbreviations

AI: adequate intake  
BLW: baby-led weaning  
BRINDA: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia  
EAR: estimated average requirement  
FFNZ: First Foods New Zealand study  
FITS: US Feeding Infants and Toddlers Study  
MSM: multiple source method  
TSF: traditional spoon feeding  
WHO: World Health Organization

Edited by T Derrick; submitted 24.03.21; this is a non-peer-reviewed article; accepted 30.03.21; published 21.04.21.

Please cite as:
Protocol

Protocol for the Pregnancy During the COVID-19 Pandemic (PdP) Study: A Longitudinal Cohort Study of Mental Health Among Pregnant Canadians During the COVID-19 Pandemic and Developmental Outcomes in Their Children

Gerald F Giesbrecht\textsuperscript{1,2,3,4}, PhD; Mercedes Bagshawe\textsuperscript{5}, BA; Melinda van Sloten\textsuperscript{1}, BA; Anna L MacKinnon\textsuperscript{2,3}, PhD; Ashley Dhillon\textsuperscript{3}, BSc; Marcel van de Wouw\textsuperscript{1}, PhD; Elnaz Vaghef-Mehrabany\textsuperscript{1}, PhD; Laura Rojas\textsuperscript{1,2}, BSc; Danielle Cattani\textsuperscript{1}, BSc; Catherine Lebel\textsuperscript{2,5*}, PhD; Lianne Tomfohr-Madsen\textsuperscript{1,2,3*}, PhD

\textsuperscript{1}Department of Pediatrics, University of Calgary, Calgary, AB, Canada
\textsuperscript{2}Alberta Children's Hospital Research Institute, Calgary, AB, Canada
\textsuperscript{3}Department of Psychology, University of Calgary, Calgary, AB, Canada
\textsuperscript{4}Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada
\textsuperscript{5}Department of Radiology, University of Calgary, Calgary, AB, Canada
*these authors contributed equally

Corresponding Author:
Gerald F Giesbrecht, PhD
Department of Pediatrics
University of Calgary
2500 University Drive
Calgary, AB, T2N 1N4
Canada
Phone: 1 403 441 8469
Email: ggiesbre@ucalgary.ca

Abstract

Background: The COVID-19 pandemic and countermeasures implemented by governments around the world have led to dramatically increased symptoms of depression and anxiety. Pregnant individuals may be particularly vulnerable to the negative psychological effects of COVID-19 public health measures because they represent a demographic that is most affected by disasters and because pregnancy itself entails significant life changes that require major psychosocial and emotional adjustments.

Objective: The PdP study was designed to investigate the associations among exposure to objective hardship caused by the pandemic, perceived stress and psychological distress in pregnant individuals, and developmental outcomes in their offspring.

Methods: The PdP study comprises a prospective longitudinal cohort of individuals who were pregnant at enrollment, with repeated follow-ups during pregnancy and the postpartum period. Participants were eligible if they were pregnant, \( \geq 17 \) years old, at \( \leq 35 \) weeks of gestation at study enrollment, living in Canada, and able to read and write in English or French. At enrollment, participants completed an initial survey that assessed demographic and socioeconomic characteristics, previous pregnancies and births, prepregnancy health, health conditions during pregnancy, medications, psychological distress, social support, and hardships experienced because of the COVID-19 pandemic (eg, lost employment or a loved one dying). For the first three months following the initial survey, participants received a monthly email link to complete a follow-up survey that asked about their experiences since the previous survey. After three months, follow-up surveys were sent every other month to reduce participant burden. For each of these surveys, participants were first asked if they were still pregnant and then routed either to the next prenatal survey or to the delivery survey. In the postpartum period, surveys were sent at 3, 6, and 12 months of infant age to assess maternal stress, psychological distress, and infant development.

Results: Participant recruitment via social media (Facebook and Instagram) began on April 5, 2020, and is ongoing. As of April 2021, more than 11,000 individuals have started the initial survey. Follow-up data collection is ongoing.

Conclusions: This longitudinal investigation seeks to elucidate the associations among hardships, maternal psychological distress, child development during the COVID-19 pandemic, and risk and resilience factors that amplify or ameliorate these
associations. The findings of this study are intended to generate knowledge about the psychological consequences of pandemics on pregnant individuals and point toward prevention and intervention targets.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/25407

(JMIR Res Protoc 2021;10(4):e25407) doi:10.2196/25407

**KEYWORDS**
pregnancy; anxiety; depression; stress; social support; resilience; COVID-19; infant development; pandemic

**Introduction**

**Background**

In December 2019, the novel SARS-CoV-2 caused an outbreak of COVID-19 in Wuhan, China, which rapidly spread around the world. COVID-19 was declared a global pandemic on March 11, 2020 [1]. Although most people with COVID-19 recover from the disease, the associated morbidity and mortality, as well as the uncertainty surrounding long-term effects, has prompted governments around the world to implement public health measures to slow and reduce the spread of COVID-19. Worldwide, these measures have included recommendations for hand and respiratory hygiene (ie, frequent handwashing; avoiding touching one’s eyes, nose, and mouth; coughing or sneezing into a bent elbow), travel restrictions, self-isolation, wearing masks in public, and physical distancing. These measures resulted in dramatic changes in the everyday life for most people, including the ways that people work, socialize, eat, and play. For example, many people were mandated to work from home, and a large number of people lost their jobs (both temporary and permanent) or saw substantial changes in their jobs. Schools and daycare centers were closed, with controversial reopening plans. Hospitals, health care facilities, and care homes limited their services and restricted visitor, caregiver, and support person access to patients and residents. Recreational facilities were closed, and public health recommendations to remain at home for anything other than essentials severely restricted people’s opportunities for recreation, physical activity, and socializing. These examples illustrate the disruption across many aspects of everyday life and the potential for uncertainty, worry, and fear that result from public health measures to limit the spread of COVID-19. A timeline of the major public health measures implemented in Canada is illustrated in Figure 1.
A Need to Understand the Effects of Pandemics on Mental Health

Although there is evidence that disasters increase symptoms and incidence of mental illness, research on the mental health consequences of epidemics and pandemics is sparse. Existing studies focus on front-line health care workers [2] and the mental health consequences of adjusting to morbidity caused by disease (eg, parents of children with congenital zika virus syndrome) [3]. However, little is known about the psychological effects of pandemic countermeasures such as quarantine, physical distancing, and shelter-in-place orders. A small study of individuals instructed to voluntarily quarantine during the 2003 severe acute respiratory syndrome (SARS) outbreak in Toronto reported elevated symptoms of posttraumatic stress disorder and depression [4]. Postdisaster studies have consistently observed increases in mental health problems following large-scale but localized events such as natural (eg, earthquakes) [5], traumatic (eg, the World Trade Center attacks) [6], and environmental (eg, Chernobyl nuclear disaster) [7] disasters.

Although most individuals will display resilience in the face of disaster, a substantial proportion will experience some psychological impairment and a smaller proportion will develop a mental health disorder [8]. The degree or severity of exposure to disaster consistently and strongly predicts greater postdisaster psychological impairment [9], with additional contributions from factors such as predisaster mental health problems, low socioeconomic status, minority ethnicity, low social support, younger age, caring for children, and personality characteristics such as neuroticism [8]. A national population-based cohort study conducted in the United Kingdom found that minority ethnicity, overweight or obesity, age over 35 years, and pre-existing health conditions increased the risk for severe mental health problems following pandemic-related events [8].

Figure 1. Timeline of COVID-19 events.
COVID-19 (eg, hospitalization required) among pregnant women [10].

There is growing concern that the public health response to the COVID-19 pandemic has created a shadow pandemic of mental illness. Public polls have repeatedly suggested widespread and dramatically elevated worries related to financial, social, and psychological well-being [11]. For example, 37% of the respondents of a nationally representative poll in the United States conducted on April 25-27, 2020, reported feeling overwhelmed from trying to work at home and balance other needs of their family [12]. A review of studies on anxiety or depression during the COVID-19 pandemic reported a prevalence of 26% and 31%, respectively [13]. Other studies have suggested that young people and women appear to be disproportionately affected by the COVID-19 pandemic [14-16].

Pregnant individuals may be particularly vulnerable to the negative psychological effects of public health measures related to COVID-19 [17], both because they represent the demographic most affected by disasters and because pregnancy itself entails significant life changes that require major psychosocial and emotional adjustments [18]. An early report from the current cohort [19] and other cohorts around the world [20-26] show that symptoms of depression and anxiety have increased dramatically among pregnant individuals during the COVID-19 pandemic, with greater fears surrounding social isolation and disease appearing to predict a greater risk of elevated symptoms.

Physical distancing policies implemented as a countermeasure to the spread of COVID-19 are especially concerning because social support buffers the negative effects of prenatal distress on both the mother and her offspring [27,28]. Social support and community cohesion are primary protective factors in the face of large-scale stressful events [29-31], and these factors may also apply to pandemics because countermeasures are known to increase a sense of social isolation [4].

**Effects of Prenatal Psychological Distress on Birth Outcomes and Child Development**

The prenatal period is a time of vulnerability for the fetus during which maternal psychological distress can have deleterious effects on fetal development. Sustained prenatal psychological distress increases the risk of prenatal and postpartum depression, prenatal infection and illness [32], miscarriage, preterm birth, and reduced birthweight [33-37]. Furthermore, children that are prenatally exposed to maternal psychological distress are more likely to have physical, behavioral, cognitive, and emotional problems than their nonexposed peers, and they are at higher risk for physical and mental health problems at a later stage [33,38-42]. Specifically, regarding disasters, a series of reports on children born to mothers exposed to the 1998 Quebec ice storm found reduced cognitive and linguistic ability [43], increased risk for obesity [44], broad changes in DNA methylation [45], and increased amygdala volume, which mediated the association between prenatal maternal stress and higher levels of externalizing behavior in these children [46]. Together, findings from disaster studies suggest that increases in psychological distress following stressful events constitute a major public health concern for physical and mental development in the generation of children prenatally exposed to the current COVID-19 pandemic.

There have also been some unexpected aspects of the COVID-19 pandemic in the context of infants. Several studies during the early pandemic [47,48], although not all [49,50], reported substantially increased rates of stillbirth and reduced rates of preterm birth and very low birthweight. The decrease in the rates of preterm births appears to be temporary, with rates increasing to their more usual levels as the pandemic wears on [51]. Pregnancy cohorts would serve an important role in identifying factors contributing to any changes in birth outcomes with significant potential to improve child development outcomes if the lessons learned can be applied to postpandemic obstetric care.

It is important to note that developmental outcomes of offspring prenatally exposed to maternal psychological distress are heterogeneous, and this heterogeneity strongly suggests that risk and resilience factors operate to increase or decrease the effects of prenatal exposures on maternal and offspring outcomes [52]. Emerging evidence during the current pandemic supports the notion that risk and resilience factors such as poverty, being a racial minority, psychological resources, and social support modulate the risk of contracting COVID-19 and for more severe psychological impairment during the pandemic [53-55]. Risk and resilience factors are likely to also operate in relation to health outcomes for children born during the pandemic. It is therefore essential that risk and resilience factors for child outcomes can be identified early to optimally direct efforts to enhance resilience and reduce risk [56].

The COVID-19 pandemic presents a novel and unprecedented opportunity to study stress and resilience in humans not only because of its worldwide scope but also because people around the globe are faced with similar hardships that result from public health countermeasures (eg, job loss, social isolation, and disrupted access to health care). The current situation replicates important features of well-established paradigms to study stress susceptibility and resilience in animal models [57], where animals exposed to the same stressor or hardship nevertheless show dramatically different behavioral, immunological, epigenetic, and neurobiological responses [58-60]. Stress-susceptible individuals exhibit considerable changes in behavioral and neurobiological responses, whereas stress-resilient individuals exhibit small or temporary changes in behavior and neurobiology. We propose that objective exposure to hardships caused by the pandemic and pandemic countermeasures constitute a major prenatal stressor and that outcomes in children will differ as a function of maternal susceptibility or resilience. Specifically, we postulate that, in general, greater objective exposure to prenatal hardship because of the pandemic will be associated with poorer maternal and infant outcomes. We also expect this effect will be moderated by maternal psychological response such that low or temporary increases in maternal distress (ie, stress resilience) will be associated with less severe outcomes compared to outcomes among mothers with similar exposure to objective hardship but with large increases in maternal distress (ie, stress susceptible).
In addition to risk and resilience factors, fetal sex and timing of exposure during gestation make significant contributions to infant outcomes. Although different effects for boys and girls and across pregnancy trimesters are commonly reported, the overall findings are heterogeneous and difficult to summarize. The most pronounced sex differences have been observed for child neural or nervous system development and temperament outcomes [61]. Timing effects likely reflect vulnerabilities to environmental input at specific points in gestation (ie, sensitive periods), suggesting that it is not possible to define the specific time at which stress exposures have the greatest effects but rather that timing effects can only be specified in terms of specific outcomes [62]. For example, several large studies on stress exposure during pregnancy found the strongest associations with poor birth outcomes and behavioral disorders when exposure occurred in the second trimester [63-65], but the strongest association with affective disorders was reported when exposure occurred in the first trimester [66], and the strongest associations with neurodevelopmental disorders were reported in the third trimester [67]. However, some exposures can have opposite effects on the same outcome depending on gestational timing; for example, lower cortisol levels in early pregnancy but higher cortisol levels in later pregnancy are associated with more optimal cognitive outcomes in infants [68]. Taken together, these findings indicate the importance of including sex and exposure timing when considering the effects of the COVID-19 pandemic on offspring outcomes.

Study Purpose
The Pregnancy During the COVID-19 Pandemic (PdP) study was designed to investigate the associations among exposure to objective hardship caused by the pandemic, perceived stress and psychological distress in pregnant individuals, and developmental outcomes in their offspring. The findings of this study are intended to provide knowledge about the psychological consequences of pandemics on pregnant individuals and their offspring and point toward prevention and intervention targets.

Methods

Primary Aims and Hypotheses

Aim 1
This study aims to determine the associations among objective exposure to hardship, perceived stress, and psychological distress among pregnant individuals during the COVID-19 pandemic. We hypothesize that greater exposure to COVID-19 stressors (eg, job loss or financial strain, death of a family member) will be associated with increased symptoms of depression, anxiety, and subjective stress.

Aim 2
This study aims to determine whether prepandemic risk factors increase vulnerability to objective COVID-19 hardship. We hypothesize that drug and alcohol use prior to pregnancy, adverse childhood experiences, minority ethnicity, low educational attainment, and poverty will increase the association between objective COVID-19 hardship and maternal psychological distress.

Aim 3
This study aims to determine whether resilience factors decrease psychological distress among pregnant individuals during the COVID-19 pandemic. We hypothesize that higher partner support, better sleep quality, and more physical activity will moderate or buffer associations between objective COVID-19 hardship and maternal psychological distress.

Aim 4
This study aims to determine the associations among exposure to prenatal objective hardship, perceived stress and psychological distress, and child development outcomes. We hypothesize that objective exposure to COVID-19 hardship will be more strongly associated with poor child development among stress-susceptible individuals (who also show high levels of subjective stress and psychological distress) than among stress-resilient individuals (who have low levels of subjective stress and psychological distress despite high exposure to objective hardship).

Secondary Aims (Hypothesis Generation)

Aim 5
This study aims to determine whether gestational age at the onset of the COVID-19 pandemic is associated with infant outcomes. It is likely that each child outcome will be more strongly associated with exposure onset at some timepoints than at other timepoints during pregnancy.

Aim 6
This study aims to identify unique features of the COVID-19 pandemic that are particularly associated with increased psychological distress. We will explore potential changes in diet, physical activity, abuse, social connection, caregiving, employment, and finances.

Aim 7
This study aims to identify unexpected aspects of the pandemic that may contribute to mental wellness or distress. We will explore potential changes in family closeness, reduction in preterm birth, and positive aspects of working from home.

Study Design and Procedures
The PdP study comprises a prospective longitudinal cohort of pregnant individuals (at enrollment) with repeated follow-ups during pregnancy and the postpartum period. Study enrollment, consent, and administration of questionnaires were conducted through Research Electronic Data Capture (REDCap) [69]. Advertisements through social media (Facebook and Instagram) directed potential participants to the study website [70] where they completed the eligibility survey and enrolled into the study. All participants signed the electronic consent form before proceeding to the first questionnaire.

An overview of the study procedure is presented in Figure 2. At enrollment, participants completed the initial survey that assessed demographic, socioeconomic, and obstetric characteristics, including age, postal code, ethnicity, household income, employment, marital status, education, country of origin, food insecurity, housing stability, history of previous
pregnancies and births, prepregnancy health, prepregnancy height and weight, current weight, health conditions prior to and during pregnancy, medications, and other measures listed below. For the first three months following the initial survey, participants received a monthly email link to complete a follow-up survey that asked about their experiences since the previous survey. After three months, the follow-up surveys were sent every other month to reduce participant burden. Thus, participants completed a maximum of five prenatal follow-up surveys in addition to the initial survey. For each of these surveys, participants were first asked if they were still pregnant, and their answer to this question routed them either to the next prenatal survey or to the delivery survey. In the postpartum period, surveys were sent at 3, 6, and 12 months of infant age. A complete list of measures and timing of data collection are presented in Table 1.

Figure 2. Study flow diagram.
### Table 1. Study measures and timepoints.

<table>
<thead>
<tr>
<th>Study measures</th>
<th>Timepoints</th>
<th>[\text{Prenatal period}]</th>
<th>[\text{Prenatal follow-up (1-5)}]</th>
<th>Delivery</th>
<th>Postnatal period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[\text{Initial survey}]</td>
<td>[\text{Prenatal follow-up (1-5)}]</td>
<td>Delivery</td>
<td>3 months</td>
<td>6 months</td>
</tr>
<tr>
<td>General measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligibility questions and consent</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic information</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health behavior (eg, substance use, physical activity, and diet)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Prior, current, and changes in medical history and conditions</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pandemic Objective Hardship Scale</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived COVID-19 threat</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distress thermometer</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edinburgh Postpartum Depression Scale (EPDS)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy-Related Anxiety Questionnaire (PRAQ/PRAQ-R2)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PROMIS(^a) Anxiety</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PROMIS Anger</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROMIS Sleep-Related Impairment</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROMIS Sleep Disturbance</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Stress Scale (PSS)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connor-Davidson Resilience Scale (CD-RISC-2)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social isolation</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Couple’s Satisfaction Index (CSI)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Support Effectiveness Questionnaire (SSEQ)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal Support Evaluation List (ISEL)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intolerance of Uncertainty Scale- Short form (IUS)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical abuse (PRAMS(^b))</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse Childhood Experiences (ACEs)(^c)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Everyday Discrimination Scale (EDS)(^c)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-compassion(^c)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intended infant feeding</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender identity and sexual orientation questions(^c)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Delivery measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery type and outcome (live birth, miscarriage, and neonatal death)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baby information and health (weight, length, sex, and NICU(^d) stay)</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth experience and COVID-19 restrictions during birth or NICU stay</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial breastfeeding questions</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COVID-19 impact on breastfeeding and bonding</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child development measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) PROMIS = Patient Report Measuring Instrument System; \(^b\) PRAMS = Parent Report Abuse Measurement System; \(^c\) ACEs = Adverse Childhood Experiences; \(^d\) NICU = Neonatal Intensive Care Unit
Study measures

<table>
<thead>
<tr>
<th>Timepoints</th>
<th>Study measures</th>
<th>Prenatal period</th>
<th>Delivery</th>
<th>Postnatal period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Initial survey</td>
<td>Prenatal follow-up (1-5)</td>
<td>3 months</td>
</tr>
<tr>
<td></td>
<td>Infant health</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Ages and Stages Questionnaire, Third Edition (ASQ-3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ages and Stages Questionnaire: Social-Emotional, Second Edition (ASQ:SE-2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brief Infant Sleep Questionnaire (BISQ)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>The Infant Behaviour Questionnaire–Revised Very Short Form (IBQ-R)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Crying patterns</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Infant feeding</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>COVID-19 disruptions to mothers’ postpartum services</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>COVID-19 disruptions to infant appointments and services</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>


| footnote b | PRAMS: Pregnancy Risk Assessment Monitoring System. |
| footnote c | These measures were only collected once. |
| footnote d | NICU: neonatal intensive care unit. |

**Study Population**

Participants included individuals who were pregnant during the COVID-19 pandemic. Participants were considered eligible if they were ≥17 years, ≤35 weeks of gestation at the time of enrollment, living in Canada, and able to read and write in English or French. There were no additional exclusion criteria. The requirement of ≤35 weeks of gestation at study enrollment was intended to allow us to collect multiple data points during a participant’s pregnancy. However, we note that, based on their due dates, some participants were at >35 weeks of gestation at initial enrollment but otherwise provided legitimate data. We plan to retain these participants for potential secondary analyses, as relevant.

Participant recruitment began on April 5, 2020, and it is currently ongoing. The recruitment goal is to obtain 9200 completed baseline surveys (see **Sample Size Considerations** below). To ensure broad representation, our advertising and surveys were available in both official Canadian languages (ie, French and English), and our social media ads target geographic regions and/or sociodemographic groups with less representation in the cohort (eg, rural communities in Northern Canada). **Figure 3** provides a graphical summary of the geographic distribution of participants across Canada enrolled in this study to date.
Study Measures

Maternal Measures

Depression Symptoms

Maternal depression symptoms experienced by the participants in the past week were assessed using the Edinburgh Postpartum Depression Scale (EPDS) [71], a self-report questionnaire with possible scores ranging from 0 to 30, where higher scores indicate more severe symptoms. A cut-off score ≥13 is used to identify individuals with clinically concerning depression symptoms [71]. Scores ≥13 during pregnancy have a sensitivity and specificity of 100% and 87%, respectively, for classifying major depression, and a positive predictive value of 33 [72]. This scale has been validated for both prenatal and postnatal assessment [73,74].

Anxiety Symptoms

General anxiety symptoms experienced by the participants in the past week were assessed using the 7-item Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety–Adult Short Form [75]. We follow the standard practice of converting raw scores to T-scores using the US general population norms; possible scores range from 36.3 to 82.7, with a mean score of 50 (SD 10). T-scores in the range of 60-69.9 are indicative of moderately elevated anxiety symptoms, and scores ≥70 are considered indicative of severely elevated anxiety symptoms [76].

Pregnancy-related anxiety symptoms (ie, worries about the health of the baby, birth, and caring for a new baby) experienced by the participants were assessed using the 10-item Pregnancy Related Anxiety Questionnaire (PRAQ) [77] in the English-language survey and the French translation of the 10-item Pregnancy Related Anxiety Questionnaire–Revised 2 (PRAQ-R2) in the French-language survey [78]. Possible scores on the PRAQ range from 10 to 40 and those on the PRAQ-R2 range from 10 to 50. Both questionnaires have acceptable internal consistency (Cronbach α=.80-.81) [78,79], and their validity is supported by extensive use in relation to both maternal and infant outcomes. There are no cut-off scores for these scales, but previous treatment studies have used a median split to define groups with higher versus lower pregnancy-related anxiety symptoms [80], with higher scores indicating more severe symptoms.

Anger

Anger experienced by the participants in the past week was assessed using the 5-item PROMIS Short Form v1.1–Anger 5a...
As with other PROMIS measures, raw scores are converted to T-scores. T-scores range from 32.9 to 82.9; the mean score was 50 (SD 10), with scores in the range of 60–69.9 indicative of moderately elevated anger and scores ≥70 indicative of severely elevated anger [76].

**Distress Thermometer**

The Distress Thermometer was used to measure the overall level of subjective distress experienced by the participants in the previous week, on a visual analog scale of 0 (“Not distressed”) to 10 (“Extremely distressed”) [81]. A cut-off score ≥4 is typically used to signify clinically concerning distress [82]. The Distress Thermometer has demonstrated good validity and temporal stability [83,84].

**Perceived Psychological Stress**

Participants’ subjective experience of psychological stress over the past month was assessed with the widely used 10-item Perceived Stress Scale (PSS) [85]. The PSS measures the degree to which participants appraise their lives as unpredictable, uncontrollable, and overloaded [86]. Scores range from 0 to 40, with higher scores indicating greater perceived stress. Reliability and validity of the PSS have been supported across multiple studies [85].

**Pandemic Objective Hardship Scale**

COVID-19 represents a novel exposure for which there are no previously developed measures. Nevertheless, previous work on objective measures of hardship resulting from exposure to natural disasters provided a principled and systematic framework for developing a new measure [87]. For instance, work by King and Laplante [87] has identified four major components of disaster exposures, which were adapted to the COVID-19 context: scope, loss, threat, and change. Scope refers to the duration and intensity of the hardship, with the former referring to the amount of time for which major aspects of participants’ lives were disrupted and the latter focusing on the number of individuals within participants’ communities who were similarly affected. Loss refers to financial, social, and physical losses experienced as a result of the pandemic. For example, the loss of employment, savings, closures of schools, and daycares represent relevant losses. Threat refers to physical and health-related consequences of exposure to the stressor. For example, being infected with SARS-CoV-2 or hospitalization of a close friend with COVID-19 represent threats to self and others. Change captures the adjustments to daily living, prenatal care, work, and social interaction caused by the COVID-19 pandemic. Relevant changes include working from home, altering a birth plan, and reductions in physical activity or diet quality. The timeframe for these items was the previous month or the previous questionnaire (ie, up to three months).

**Prenatal and Postnatal Care**

Changes and disruptions to prenatal and postnatal care received by the participants were assessed using a series of questions tailored to our study relating to the way that prenatal care was delivered, cancellations, ability to bring partner to appointments, and changes to birth plans. Participants are also asked to evaluate the impact of these changes on the quality of care they and their baby have received in the past three months.

**Perceived COVID-19 Threat**

The degree to which participants feel that COVID-19 was a threat to their health or the health of their baby at any time during the pandemic was assessed by three items developed for the study: (1) “How much do (did) you think your life is (was) in danger during the COVID-19 pandemic?” (2) “How much do (did) you think your baby’s life is (was) in danger at any time during the COVID-19 pandemic?” and (3) “How much are you worried that exposure to the COVID-19 virus will harm your baby?” Responses were scored on a 100-point sliding scale, with the left anchor indicating 0 points (“Not at all”); the middle anchor, 50 points (“Somewhat”); and the right anchor, 100 points (“Very much so”).

**Adverse Childhood Experiences**

Adverse childhood experiences (ACEs) of the participants were assessed using a 10-item measure [88] of early-life adversity (age 0–18 years) across three domains: abuse (emotional, physical, and sexual), neglect (physical and emotional), and exposure to household dysfunction (domestic violence, substance abuse, mental illness, parental separation or divorce, and incarcerated household member). The ACEs questionnaire is widely used and has demonstrated good reliability and internal consistency (Cronbach α=.81) [89], as well as adequate test-retest reliability (weighted κ=0.64) [90]. The occurrence of individual ACEs is summed to create the ACE score with a potential range of 0–10. Based on previous work that has examined dose-response relationships, participants were coded as having experienced 0, 1, 2, 3, or ≥4 ACEs [88,91,92].

**Perceived Social Support**

Participants’ perception of the quality of social support received from their romantic partner over the previous three months was assessed using the 35-item Social Support Effectiveness Questionnaire (SSEQ) [79]. Within the domains of emotional, informational, and task support, participants were asked to rate their experience over the past three months on a 5-point scale: (1) how well the quantity of support received from their partner matched the amount they wanted, (2) whether they wished the support had been different somehow, (3) how skillful their partner was at providing support, (4) how often it was difficult to solicit support, and (5) whether their partner offered support without being asked. Participants also rated the extent to which they perceived their partner’s support as negatively infringing upon their self-efficacy or self-esteem. Internal consistency of the SSEQ is strong (Cronbach α=.87), and it has previously been used to distinguish levels of social support in samples of pregnant individuals [79,93,94]. Total scores can range from 0 to 80, with higher scores indicating more effective support.

Perceived social support was also assessed using the 12-item Interpersonal Support Evaluation List (ISEL) [95,96], to determine appraisal (eg, advice or problem solving), belongingness (eg, shared experiences), and tangible support (eg, help with daily chores) over the past three months. Here, we focus on the total score, which is the sum of the three subscales. Scores range from 12 to 48, with a higher score indicating greater perceived support.
Social Isolation
Feelings of social isolation were assessed using the following item: “During the COVID-19 pandemic, I have felt more alone than usual.” Participants’ responses were on a 100-point sliding scale, with the left anchor indicating 0 points (“Not at all”); the middle anchor, 50 points (“Somewhat”); and the right anchor, 100 points (“Very much so”).

Relationship Quality
Participants’ perception of partner relationship quality was assessed using the 4-item Couple Satisfaction Index (CSI-4) [97]. Scores range from 0 to 21, with higher scores indicating higher levels of relationship satisfaction. CSI-4 scores below 13.5 suggest notable relationship dissatisfaction. Changes in relationship quality as a function of the pandemic were assessed for relationships with partner, children, and close friends and family by using items generated for the study. For each relationship type, participants were asked how the COVID-19 pandemic affected their relationship. Responses were recorded on a 100-point sliding scale, with the left anchor indicating 0 points (“It has strained our relationship”); the middle anchor, 50 points (“Not much has changed”); and the right anchor, 100 points (“It has brought us closer together”).

Physical Abuse
Experiences of physical harm in the 12 months prior to pregnancy and during the current pregnancy were assessed using two items adapted from the Pregnancy Risk Assessment Monitoring System (PRAMS) [98]. These items were queried with regard to the 12-month periods before pregnancy, since the last survey, and after giving birth.

Physical Activity
Participants reported their physical activity using a modified form of the Godin-Shephard Leisure-Time Exercise Questionnaire (GLTEQ) in a typical week over the past month [99]. Participants reported the number of days per week in which they engaged in mild (eg, light walking), moderate (eg, brisk walking), and strenuous (eg, running) exercise of more than 15 minutes. Definitions of each category of physical activity were provided. The total score was calculated per standard scoring procedure for GLTEQ, by multiplying episodes of mild exercise by 3, moderate exercise by 5, and strenuous exercise by 9. Participants with scores below 14 are considered sedentary, those with scores ranging from 14 to 23 are considered moderately active, and those with scores equal to or more than 24 are considered active. An additional item was included asking participants how their level of physical activity changed because of the COVID-19 pandemic. Responses were recorded on a 5-point Likert-type scale with the following options: “Substantially decreased” (1 point), “Somewhat decreased” (2 points), “No change” (3 points), “Somewhat increased” (4 points), and “Substantially increased” (5 points).

Sleep Quality
Sleep disturbance and impairment in the past week were assessed using the 4-item PROMIS Sleep Disturbance–Short Form 4a and the 4-item PROMIS Sleep-Related Impairment–Short Form 4a [100]. As with other PROMIS measures, raw scores are converted to T-scores. T-scores ranging from 60 to 69.9 are considered moderate problems, and scores ≥70 are considered severely elevated. Sleep duration (ie, hours of sleep per night) was assessed with a single item from the Pittsburgh Sleep Quality Index [101].

Diet
Changes in diet and eating patterns were assessed using a questionnaire developed for the study to determine how participants’ diet during the COVID-19 pandemic differed from their prepandemic diet and the reasons for the change. Dietary changes were assessed in the following categories: fresh vegetables or fruits, dairy, meats, canned or dried foods, fast foods, take-out or home delivery, and sweets or snacks. For each category, participants responded on a 5-point Likert-type scale, with the following options: “I eat much more” (1 point), “I eat more” (2 points), “I eat about the same” (3 points), “I eat less” (4 points), and “I eat much less” (5 points). If participants reported a change in their diet (ie, did not select “I eat about the same” option), then they were asked about the reason for the change with the following response options: “Can no longer afford,” “Can’t go grocery shopping frequently,” “Can spend more time cooking and preparing food,” “Change in craving,” and “Other (specify below)” (scores: Yes=1, No=0). Although changes in craving were not expected as a direct result of the pandemic, they are commonly reported during pregnancy [102]; it was therefore important to disambiguate such pregnancy-related changes in diet from changes that are more directly related to the pandemic.

Substance Use
Substance use prior to (ie, in the 12 months before pregnancy) and during the current pregnancy (ie, in the past month) were assessed using a self-report measure separately for alcohol, cannabis, tobacco, and illicit drugs. Participants were asked how many days per week they consumed each substance and how many drinks or products per day they typically consumed. Participants were also asked whether they changed their use patterns in pregnancy and if they had, they were asked when they changed their use patterns.

Coping and Resilience
Participants’ perceived ability to cope with stressful situations was assessed with the 2-item Connor-Davidson Resilience Scale (CD-RISC 2), which demonstrates good test-retest reliability and convergent and divergent validity [103]. Each item is rated on a 5-point Likert scale ranging from 0 (“Not true at all”) to 4 (“True all the time”); total scores can range from 0 to 8, with higher scores indicating more successful coping. We also included an open-ended question about coping: “People are responding to the pandemic in many ways. Can you tell us what things you are doing to cope with the COVID-19 pandemic?”

The 27-item Intolerance of Uncertainty Scale (IUS) was used to measure the extent to which participants believe uncertainty is stressful, upsetting, negative, unfair, and leads to the inability to take action [104,105]. Items are rated on a scale ranging from 1 (“Not at all characteristic of me”) to 5 (“Entirely characteristic of me”). Previous research shows that the IUS has excellent internal consistency (Cronbach α=91) and good test-retest
reliability ($r=0.74$), and it is highly correlated with symptoms of generalized anxiety disorder [104].

Self-compassion, a trait level form of resilience, was assessed using the 12-item short form of the Self-Compassion Scale–Short Form (SCS-SF) [106]. The SCS measures the tendency to treat oneself with kindness and understanding rather than harsh self-judgment, to recognize imperfections and suffering as part of the human experience, and to have mindful awareness and tolerance of negative thoughts and feelings. Each item is rated on a 5-point Likert scale ranging from 1 (“Almost never”) to 5 (“Almost always”); total scores range from 10 to 60, with higher scores indicating more self-compassion. The SCS has adequate internal consistency (Cronbach $\alpha=0.86$) and strong correlation ($r=0.97$) with the long form [106].

**Discrimination**

Participants’ experiences of discrimination were assessed using the 5-item Everyday Discrimination Scale [107]. This questionnaire asks about day-to-day experiences of discrimination, including being treated differently than other people and feeling threatened or harassed. Responses items ranged from 0 (“Never”) to 6 (“Almost every day”). For items rated as more than “A few times a year,” participants were also asked a follow-up question about what they think is the main reason for these experiences.

**Vaccination**

Participants were asked if they planned to receive a COVID-19 vaccine, if they had received a COVID-19 vaccine (which vaccine, if yes), and the dates of the doses (or the gestational age if they were still pregnant).

**Child Development Measures**

Parent report of child outcomes were assessed at 3, 6, and 12 months postpartum.

**Developmental Milestones**

The Ages and Stages Questionnaire, Third Edition (ASQ-3) [108] is a widely used parent-reported and norm-referenced developmental screening tool [109] that assesses delays in child development across five domains: communication, gross motor, fine motor, problem-solving, and personal adaptive skills. The ASQ-3 has been identified by the American Academy of Pediatrics as a high-quality tool for use in clinical practice to screen for delayed developmental milestones in children [110]. Sensitivity and specificity of the ASQ-3 are both 86% for distinguishing between children at risk for developmental delay and children not at risk. Parents rate each of the 30 items on a scale ranging from “Yes” (10 points), “Sometimes” (5 points), or “Not yet” (0 point) based on the infant’s current ability.

**Socioemotional Development**

The Ages and Stages Questionnaire: Social-Emotional, Second Edition (ASQ-SE-2) [111] is a validated and widely used parent-report screening tool in 7 areas of socioemotional development: self-regulation, compliance, social communication, adaptive functioning, autonomy, affect, and interaction with people. A total score is calculated to index overall socioemotional problems. Parents rate each of the 27 items on a scale ranging from “Often or always” (0 point), “Sometimes” (5 points), or “Rarely or never” (10 points) based on the infant’s usual behavior.

**Temperament**

The Infant Behavior Questionnaire–Revised Very Short Form (IBQ-R) [112,113] is a 37-item parent-report measure of infant temperament. Three broad dimensions of temperament are assessed: negative emotionality, regulatory capacity/orienting, and positive affectivity/surgency. Parents report their observations of specific infant behaviors in the past week using a 7-point Likert scale ranging from “Never” to “Always.” Each item also had a “Does not apply” option. Scores on each dimension can range from 1 to 7, with higher scores reflecting stronger evidence of each dimension. The IBQ-R has strong psychometric properties and is widely used in the child development literature [114,115].

**Crying**

Periods of persistent infant crying (defined as half an hour or more during which the baby would not settle) in the past week were assessed using several items from the Crying Patterns Questionnaire [116]. The validity of the Crying Patterns Questionnaire, relative to a cry diary, has been supported [117].

**Infant’s Sleep Quality**

Parents’ report of their infant’s sleep quality in the past two weeks was assessed using the 19-item Brief Infant Sleep Questionnaire (BISQ)–Revised Short Form [118]. In addition to a total score, three subscale scores are calculated: infant sleep (5 items), parent perception (3 items), and parent sleep-related behaviors (11 items). The measure is scored using an age-based and norm-referenced system [119]. The total score and each subscale score are scaled from 0 to 100, with higher scores indicating better sleep quality, more positive perception of infant sleep, and parent behaviors that promote healthy and independent sleep. The BISQ is widely used, and its reliability and validity have been documented [120,121].

**Feeding**

Based on previously published questionnaires [122,123], we assessed onset, duration, and proportion of breastfeeding and formula feeding by using a series of maternal-report questions that conform to the World Health Organization’s breastfeeding categories: exclusive breastfeeding, predominant breastfeeding, mixed feeding, and bottle feeding [124].

**Sample Size Considerations**

To power the study adequately for each of the hypotheses, we conducted a sample size calculation for Hypothesis 5, which will test the timing effects of prenatal exposure on child development outcomes. We choose this hypothesis because it requires the largest sample, and adequately powering this hypothesis will also adequately power the other hypotheses. We used G*Power (version 3.1) [125] to estimate the sample size required to test differences in the proportion of children not meeting developmental milestones. About 26% of children in Canada did not meet developmental milestones in one or more area of development prior to the COVID-19 pandemic [126]. To detect an increase in proportion of 0.09 not meeting developmental milestones (which is equal to the interprovincial
variability in the proportion of children not meeting developmental milestones), assuming a power of 0.90 and α=0.01, a sample size of 2184 is required. To conduct adequately powered analyses stratified by trimester, we require a sample of 6552 with infant milestone data. Allowing for a 10% attrition due to miscarriage and stillbirth and an additional 30% attrition due to other reasons after the initial survey, we plan to recruit a total sample of 9200.

Data Analysis
Data visualization and screening will be conducted to determine what, if any, data manipulations are required. Regression-based analysis are planned to address study hypotheses. Logistic regression will be used with categorical outcomes, and multivariable regression will be used for continuous outcomes. Longitudinal analyses will be conducted using multilevel modeling. All analyses will include covariates deemed important to control for confounding and to increase the precision of the model. Planned subgroup analyses include grouping based on established cut-off scores for measures of anxiety and depression symptoms, analyses comparing individuals with confirmed COVID-19 to those who did not have COVID-19, and analyses by child sex for child outcomes. We also plan to examine time-related factors, including timing effects related to trimester of pregnancy when the pandemic was declared and the influence of pandemic phase (ie, how stress, distress, and fear change over time) on outcomes. Missing data will be assessed to determine what treatments are required. Given the risk of attrition bias in longitudinal studies, we will use a missing data strategy that yields unbiased estimates (eg, maximum likelihood and multiple imputation).

Ethical Considerations
This study received ethics approval (REB20-0500) from the University of Calgary Conjoint Health Research Ethics Board on March 26, 2020. All participants were required to voluntarily agree to participate in this study and sign the electronic informed consent form prior to providing any data.

Data Management and Availability
Study data were collected and managed using REDCap electronic data capture tools hosted at the University of Calgary and the University of Alberta [69]. REDCap is a secure, web-based application designed to support data capture for research studies, providing (1) an intuitive interface for validated data entry, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for importing data from external sources.

Metadata will be included in the Canadian Research Advancement through Cohort Cataloguing and Harmonization (REACH) project [127].

Data are available upon reasonable request made to the corresponding author.

Results
Participant recruitment via social media (Facebook and Instagram) began on April 5, 2020, and is ongoing. As of April 2021, more than 11,000 individuals started the initial survey. Follow-up data collection is ongoing.

Discussion
The design and implementation of this study protocol were executed during the early phase of the pandemic with data collection starting on April 5, 2020. Because of the evolving nature of the pandemic, some of the questions specific to the pandemic required modification, and additional questions were added to reflect the emerging issues. For example, beginning in June 2020, when provincial governments began to implement phased approaches to relaunching the economy, we added questions about perceptions of these relaunch plans.

Strengths of this study include its prospective longitudinal design, implementation at an early phase of the pandemic, sample size, recruitment (and representation) from every province and territory in Canada, a measure of objective exposure to pandemic hardships, use of measures with strong psychometric properties, and measurement of many potential confounding variables. Limitations include reliance on self-report measures that are not diagnostic in nature, the potential to attract participants with higher levels of psychological distress, the potential for attrition bias because of differential loss to follow-up, and the use of a cohort design that limits causal interpretation.

This longitudinal investigation seeks to elucidate the associations between hardships caused by the COVID-19 pandemic, maternal psychological distress, child development, and risk and resilience factors that amplify or ameliorate these associations. The findings of this study are intended to provide knowledge about the psychological consequences of pandemics on pregnant individuals and point toward prevention and intervention targets.

Acknowledgments
The authors thank the participants of the Pregnancy During the COVID-19 Pandemic study for their time and effort. The authors are grateful to Martina Min for assistance with the creation of Figure 2, Tasha Christie for assistance with Table 1, and Angélica Boucher for assistance with French-language translation. This project is supported by funds from the Owerko Center at the Alberta Children’s Hospital Research Institute. LTM was supported by a Career Development Award from the Canadian Child Health Clinician Scientist Program. MW is supported by an Eyes High Postdoctoral Fellowship from the University of Calgary. LR is supported by a graduate fellowship from the Alberta Children’s Hospital Research Institute. ALM is supported by postdoctoral fellowships from the Social Sciences and Humanities Research Council (SSHRC) and ACHRI.
Authors' Contributions
GFG, CL, and LTM conceptualized the study, obtained funding, and supervised data collection. GFG wrote the first draft of the manuscript. MB, MS, ALM, AD, MW, EVM, LR, and DC assisted with study design and data collection. All authors revised and approved the final manuscript.

Conflicts of Interest
None declared.

References


70. Pregnancy During the Pandemic. URL: https://www.pregnancyduringthepandemic.com/ [accessed 2021-04-16]


89. Bruskas D, Tessin DH. Adverse childhood experiences and psychosocial well-being of women who were in foster care as children. Perm J 2013;17(3):e131-e141 [FREE Full text] [doi: 10.7812/TPP/12-121] [Medline: 24355905]


98. PRAMS Questionnaires. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/prams/questionnaire.htm [accessed 2021-04-16]


https://www.researchprotocols.org/2021/4/e25407


Abbreviations

ACEs: Adverse Childhood Experiences
ASQ-3: Ages and Stages Questionnaire, Third Edition
BISQ-R: Brief Infant Sleep Questionnaire–Revised
CD-RISC 2: Connor-Davidson Resilience Scale
CSI-4: 4-item Couple Satisfaction Index
EDS: Everyday Discrimination Scale
EPDS: Edinburgh Postnatal Depression Scale
GLTEQ: Godin-Shephard Leisure-Time Exercise Questionnaire
IBQ-R: Infant Behavior Questionnaire–Revised
ISEL: Interpersonal Support Evaluation List
IUS: Intolerance of Uncertainty Scale
PdP: Pregnancy During the COVID-19 Pandemic
PRAMS: Pregnancy Risk Assessment Monitoring System
PRAQ: Pregnancy Related Anxiety Questionnaire (English)
PRAQ-R2: Pregnancy Related Anxiety Questionnaire–Revised 2 (French)
PROMIS: Patient-Reported Outcomes Measurement Information System
PSS: Perceived Stress Scale
REDCap: Research Electronic Data Capture
SARS: severe acute respiratory syndrome
SCS-SF: Self-Compassion Scale–Short Form
SSEQ: Social Support Effectiveness Questionnaire

Edited by G Eysenbach; submitted 31.10.20; peer-reviewed by M Davenport, A Falise, S Ge; comments to author 02.04.21; revised version received 08.04.21; accepted 13.04.21; published 28.04.21.

Please cite as:
Protocol for the Pregnancy During the COVID-19 Pandemic (PdP) Study: A Longitudinal Cohort Study of Mental Health Among Pregnant Canadians During the COVID-19 Pandemic and Developmental Outcomes in Their Children
JMIR Res Protoc 2021;10(4):e25407
URL: https://www.researchprotocols.org/2021/4/e25407
doi:10.2196/25407
PMID:33848971

©Gerald F Giesbrecht, Mercedes Bagshawe, Melinda van Sloten, Anna L MacKinnon, Ashley Dhillon, Marcel van de Wouw, Elnaz Vaghef-Mehrabany, Laura Rojas, Danielle Cattani, Catherine Lebel, Lianne Tomfohr-Madsen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.
Protocol

Development of an Emergency Department–Based Intervention to Expand Access to Medications for Opioid Use Disorder in a Medicaid Nonexpansion Setting: Protocol for Engagement and Community Collaboration

Lauren A Walter¹*, MD; Li Li²*, MD; Joel B Rodgers¹, MA; Jennifer J Hess³, MD; Rachel M Skains¹, MD; Matthew C Delaney¹, MD; James Booth¹, MD; Erik P Hess³*, MD, MSc

¹Department of Emergency Medicine, University of Alabama at Birmingham, Birmingham, AL, United States
²Department of Psychiatry and Behavioral Neurobiology, University of Alabama at Birmingham, Birmingham, AL, United States
³Department of Emergency Medicine, Vanderbilt University, Nashville, TN, United States
* these authors contributed equally

Corresponding Author:
Lauren A Walter, MD
Department of Emergency Medicine
University of Alabama at Birmingham
619 19th St S
OHB 251
Birmingham, AL, 35249
United States
Phone: 1 205 975 9358
Email: lwalter@uabmc.edu

Abstract

Background: The opioid epidemic has disproportionately impacted areas in the Appalachian region of the United States. Characterized by persistent Medicaid nonexpansion, higher poverty rates, and health care access challenges, populations residing in these areas of the United States have experienced higher opioid overdose death rates than those in other parts of the country. Jefferson County, Alabama, located in Southern Appalachia, has been especially affected, with overdose rates over 2 times greater than the statewide average (48.8 vs 19.9 overdoses per 10,000 persons). Emergency departments (EDs) have been recognized as a major health care source for persons with opioid use disorder (OUD). A program to initiate medications for OUD in the ED has been shown to be effective in treatment retention. Likewise, continued patient engagement in a recovery or treatment program after ED discharge has been shown to be efficient for long-term treatment success.

Objective: This protocol outlines a framework for ED-initiated medications for OUD in a resource-limited region of the United States; the study will be made possible through community partnerships with referral resources for definitive OUD care.

Methods: When a patient presents to the ED with symptoms of opioid withdrawal, nonfatal opioid overdose, or requesting opioid detoxification, clinicians will consider the diagnosis of OUD using the Diagnostic and Statistical Manual of Mental Disorders (fifth edition) criteria. All patients meeting the diagnostic criteria for moderate to severe OUD will be further engaged and assessed for study eligibility. Recruited subjects will be evaluated for signs and symptoms of withdrawal, treated with buprenorphine-naloxone as appropriate, and given a prescription for take-home induction along with an intranasal naloxone kit. At the time of ED discharge, a peer navigator from a local substance use coordinating center will be engaged to facilitate patient referral to a regional substance abuse coordinating center for longitudinal addiction treatment.

Results: This project is currently ongoing; it received funding in February 2019 and was approved by the institutional review board of the University of Alabama at Birmingham in June 2019. Data collection began on July 7, 2019, with a projected end date in February 2022. In total, 79 subjects have been enrolled to date. Results will be published in the summer of 2022.

Conclusions: ED recognition of OUD accompanied by buprenorphine-naloxone induction and referral for subsequent long-term treatment engagement have been shown to be components of an effective strategy for addressing the ongoing opioid crisis. Establishing community and local partnerships, particularly in resource-limited areas, is crucial for the continuity of addiction care and rehabilitation outcomes.
opiod use disorder; mediation for opioid use disorder; emergency medicine; buprenorphine; peer support services

Introduction

Background

The Appalachian region of the Eastern United States has been disproportionately affected by the opioid epidemic. In 2017, the death rate for opioid overdoses in Appalachian counties was 72% higher than that in non-Appalachian counties [1]. The number of opioid-related deaths in Appalachia has now surpassed the number of deaths from motor vehicle accidents [2]. A Medicaid nonexpansion state, Alabama, is located in Southern Appalachia and has been particularly affected by the opioid epidemic. Since 2014, Alabama has led the nation with the highest rate of opioid prescriptions (107.2 prescriptions for every 100 persons), nearly two-fold greater than the national average [3]. Following efforts implemented by the US Drug Enforcement Administration (DEA) in 2010 to decrease opioid prescription rates, the demand for more potent opioids such as heroin and fentanyl increased, resulting in an unintended increase in the number of overdose deaths nationally [4]. This trend also occurred in Alabama, where there was a significant (11.1%) increase in the age-adjusted rate of drug overdose deaths from 2016 to 2017 [3]. Jefferson County in Alabama is a hot spot for opioid use disorder (OUD) and its complications. An analysis of naloxone administration in 2016 revealed that Jefferson County had 48.8 overdoses per 10,000 persons, a rate significantly higher than the statewide average of 19.9 per 10,000 persons [5]. In 2018, the Jefferson County Coroner’s office confirmed 228 illicit drug deaths, 171 of which were because of an opioid overdose [6].

Patients with OUD often present to emergency departments (EDs) to treat opioid-related conditions, including nonfatal opioid overdose [7]. Persons with OUD are high users of ED services. People who inject drugs have been shown to use EDs over 3 times more frequently than the general population [8]. Mortality after ED discharge following nonfatal opioid overdose has been shown to be as high as 5.5% within 1 year, highlighting a vulnerable population that may benefit from ED intervention [9]. Beyond acute medical stabilization, ED interventions for this patient population have traditionally been limited to brief cessation counseling and referral for drug rehabilitation. Recently, a brief negotiation interview (BNI), the treatment of withdrawal with buprenorphine-naloxone as appropriate, and referral to treatment have been shown to be an effective strategy for OUD treatment. In a randomized trial comparing referral alone with BNI plus referral and BNI plus referral plus buprenorphine, ED patients randomized to the BNI plus referral plus buprenorphine arm had a nearly two-fold higher rate of retention in addiction treatment and significantly lower use of inpatient addiction services at 30 days [10]. Despite this evidence, the widespread adoption has lagged. The implementation gap is particularly marked in Alabama, where, according to the Substance Abuse and Mental Health Services Administration (SAMHSA) database and other publicly available data, only 2.18% (138/6342) of physicians had completed the Drug Addiction Treatment Act (DATA) 2000 waiver training, a requirement to prescribe medications for opioid use disorder (MOUD) such as buprenorphine [11,12]. Data elsewhere suggest that an even smaller percentage of X-waivered clinicians actively use their X-waiver to provide MOUD [13]. The shortage of clinicians providing MOUD in Alabama is profound.

Similarly, although the ED has long been recognized as the health care system’s “front door,” providing critical access and referral to primary care and specialty services, it often does not translate to effective linkage to substance use disorder (SUD) treatment. In Appalachia, an insufficient supply of behavioral and public health services targeting opioid misuse contributes to higher rates of opioid misuse and mortality in the region [2]. Engagement in long-term rehabilitation and MOUD, however, are an integral part of the recovery process for patients with OUD. Patients with SUD who are directly admitted to a treatment facility have been found to be 30 times more likely to enroll than those who were indirectly referred, as is most often the case in the ED setting [14]. An additional important consideration is the resource-limited setting in which many EDs in non-Medicaid expansion states operate, particularly when considering treatment options for low-income, uninsured patients. Therefore, developing community partnerships that facilitate linkage to care is a critical component of OUD treatment after discharge from the ED.

Objective

The overall objective of this implementation project is to increase the number of persons with OUD who present to either of 2 academic EDs in Jefferson County to receive MOUD and community treatment referral, thereby improving OUD outcomes in this region. We will accomplish this objective by initiating buprenorphine-naloxone in the ED, activating bedside referral services, and linking patients to longitudinal addiction care. Table 1 outlines the goals and objectives of this SAMHSA-funded public health intervention (SAMHSA Award #1H79TI081609). As the first ED-initiated medication-assisted treatment (MAT) program in Alabama, a resource-limited, non-Medicaid expansion state, we aim to demonstrate the feasibility and effectiveness of an ED-MOUD program in this region and identify barriers and potential solutions to expand access to MOUD further. We hope our experience will inform other resource-limited regions in the United States that may be seeking to improve access to MOUD and treatment options for persons with OUD.
Table 1. Goals and objectives for the Emergency Department-Medication for Opioid Use Disorder Therapy Demonstration Project in Jefferson County, Alabama.

<table>
<thead>
<tr>
<th>Goals for Jefferson County, Alabama</th>
<th>Measurable objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase the number of clinicians with DATA(^a) 2000 training to identify and treat patients with OUD(^b)</td>
<td>Obtain DATA 2000 waiver training in at least 75% of ED(^c) attending and licensed resident physicians by February 28, 2020. (25 UAB-ED(^d) MDs(^e) by July 2, 2019; an additional 30 MDs by February 28, 2020).</td>
</tr>
<tr>
<td>Develop and implement an electronic evidence-based induction and referral to treatment protocol initiated in the ED</td>
<td>Implement an electronic evidence-based order set to standardize OUD diagnosis, determine the need for buprenorphine induction, activate bedside referral services, and initiate outpatient MOUD(^f) in the University and Highlands EDs by July 2, 2019.</td>
</tr>
<tr>
<td>Operationalize the protocol to increase the number of individuals with OUD referred for MOUD</td>
<td>Treat and refer 272 individuals with OUD for MOUD over the 3-year project period (56 in year 1+108 in year 2+108 in year 3).</td>
</tr>
<tr>
<td>Improve retention in care for individuals who have been diagnosed with OUD</td>
<td>Increase treatment completion rates by individuals with OUD throughout the program by 5%-10% each year (year 1=baseline).</td>
</tr>
<tr>
<td>Decrease opioid overdose–related deaths</td>
<td>In collaboration with other community-based initiatives and public health interventions, decrease opioid overdose–related deaths in Jefferson County by 30% over 3 years.</td>
</tr>
</tbody>
</table>

\(^a\)DATA: Drug Addiction Treatment Act.  
\(^b\)OUD: opioid use disorder.  
\(^c\)ED: emergency department.  
\(^d\)UAB-ED: University of Alabama at Birmingham Hospital Emergency Department.  
\(^e\)MD: Medicine Doctor.  
\(^f\)MOUD: medications for opioid use disorder.

Methods

Study Population

The population for this project includes patients seeking treatment for OUD who present to either of 2 University of Alabama at Birmingham (UAB) Hospital EDs located in Jefferson County. The University Hospital ED is a tertiary care academic ED with >70,000 annual patient visits in 2018, and the Highlands Hospital ED is an urban community ED with 30,000 annual patient visits in the same year. To estimate the number of persons potentially impacted by this public health intervention, we conducted an *International Classification of Diseases, Tenth Revision*, search of our electronic health record (EHR) from January 1, 2018, to December 31, 2018 [15]. We identified 2395 unique patients who presented to the ED to treat nonfatal opioid overdose, opioid withdrawal, or opioid detoxification. This prevalence rate, approximately 240 per 10,000 people, is above the national average of opioid-related ED visits [16]. The mean age was 39.8 years (SD 12.4 years), and most patients were male (1432/2395, 59.79%) and White (1861/2395, 77.70%); 50.02% (1198/2395) of the patients were uninsured (Table 2).

The geographic catchment area for this initiative, Jefferson County, is the largest and most populous county in the state of Alabama, encompassing the city of Birmingham and 29 additional municipalities. Jefferson County has a total area of 1124 square miles and a population of 659,300, of which 44% are African American, 53% Caucasian, and 17% live below the federal poverty level [17]. Admitted patients will be excluded from this study. Variability in inpatient treatment plans and logistical inability to

Table 2. Baseline characteristics of emergency department patients seen for opioid overdose, opioid withdrawal, or seeking detoxification in 2018 (n=2395).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>39.8 (12.4)</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>1433 (59.83)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>1860 (77.66)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>485 (20.25)</td>
</tr>
<tr>
<td>Other</td>
<td>50 (2.09)</td>
</tr>
<tr>
<td><strong>Insurance status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Privately insured</td>
<td>380 (15.87)</td>
</tr>
<tr>
<td>Publicly insured (Medicare or Medicaid)</td>
<td>817 (34.11)</td>
</tr>
<tr>
<td>Self-pay or uninsured</td>
<td>1198 (50.02)</td>
</tr>
</tbody>
</table>

https://www.researchprotocols.org/2021/4/e18734  
JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e18734 | p.425  
(page number not for citation purposes)
consistently provide a warm handoff to a community partner (see Community Partnerships section below) make enrollment of admitted patients less feasible in our setting. Other vulnerable populations, including prisoners, minors, pregnant patients, or those unable to consent, will also be excluded.

**Implementation Plan**

Enrollment will occur in the ED 24 hours per day, 7 days per week. To be feasible in our context, the protocol developed by D’Onofrio et al [10] will be modified in the following ways: first, given the large number of patients who present to UAB EDs with a primary complaint of nonfatal opioid overdose, opioid withdrawal, or requesting opioid detoxification, these chief complaints will serve to identify patients with suspected OUD rather than a universal screening method. Emergency physicians’ primary role in this project is to identify patients with OUD who may be eligible for ED-initiated MOUD and referral to treatment. If a patient meets 4 or more Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) (DSM-5) criteria (moderate to severe OUD), physicians will be asked to conduct a BNI to explore the individual’s motivation to engage in treatment (see the Physician Education and Engagement section; DATA 2000 waiver training plus supplemental physician education encompasses DSM-5 OUD criteria and BNI training). The BNI typically takes ≤5 to 15 minutes [18]. If the patient appears motivated, the physician will activate an order set within the EHR to notify the research staff in real time. Research staff will assess the patient for eligibility, conduct enrollment, and assist with linkage to care via direct communication with community referral partners. Physicians may provide a dose of buprenorphine-naloxone in the ED if the patient is in active opioid withdrawal (may be repeated if necessary for ongoing withdrawal symptoms) and will be asked to provide a buprenorphine-naloxone bridge prescription at the time of patient discharge from the ED [10]. Patients provided with a buprenorphine-naloxone prescription will also be provided with in-person (from the provider) and handout instructions regarding buprenorphine induction and titration. Patients who are not eligible or who do not wish to enroll will be provided the appropriate existing standard of care, including a list of referral resources for OUD for self-access at the time of ED discharge.

It must be noted that the prescription drug monitoring program (PDMP), a statewide controlled substance prescription database, is seamlessly integrated into the EHR at UAB and is now the standard of care in our hospitals. The PDMP for each patient is visible to all providers registered with the Alabama PDMP without the need to access an external website. A review of PDMP records before initiation of MOUD is an approach recommended by a recent American College of Emergency Physicians policy statement [19]. The PDMP will be consistently used for this study to (1) assist providers in the diagnosis of OUD, when applicable (eg, multiple overlapping opioid prescriptions from multiple providers), and (2) thoughtfully consider the treatment plan of patients who may return to the ED for MOUD refill prescriptions (see the Plans to Mitigate Risk of Diversion section).

**Clinical Protocol**

To alert research staff to potential study participants, we developed a custom OUD order set in Cerner that delivers automated, real-time, electronic notifications to research personnel when activated by an emergency physician (Table 3). A departmental protocol guides the initiation of the OUD order set by physicians at any time a patient meets the DSM-5 criteria for at least moderate OUD. Following the electronic alert, trained research staff, consisting of a research coordinator and/or research assistant, collaborate directly with the ordering the clinician and the patient to ascertain appropriateness for study inclusion. Research staff will assess patients for enrollment eligibility, complete the enrollment process, and facilitate handoff to community referral partners. If enrolled, research staff will also contact peer navigators to be dispatched to the bedside by the Recovery Resource Center (RRC)—a Jefferson County community substance abuse treatment coordinating center—to facilitate linkage to outpatient addiction treatment.
Table 3. Custom opioid use disorder electronic health record order set.

<table>
<thead>
<tr>
<th>UED&lt;sup&gt;a&lt;/sup&gt; opiate use disorder (initiated pending)</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient care</strong></td>
<td></td>
</tr>
<tr>
<td>Communication order nursing</td>
<td>• Complete COWS&lt;sup&gt;b&lt;/sup&gt; assessment</td>
</tr>
<tr>
<td>Communication order nursing</td>
<td>• Verify the patient’s contact information for follow-up purposes</td>
</tr>
<tr>
<td><strong>Medications</strong></td>
<td></td>
</tr>
<tr>
<td><strong>COWS 8-13</strong></td>
<td></td>
</tr>
<tr>
<td>Buprenorphine-naloxone (buprenorphine [dosed with naloxone])</td>
<td>• 4 mg, Tab-SL&lt;sup&gt;c&lt;/sup&gt;, sublingual, every 1 hour; now, PRN&lt;sup&gt;d&lt;/sup&gt;, other (see comment below regarding observation prior to second dose), 2 doses</td>
</tr>
<tr>
<td></td>
<td>• Administer 4 mg now. Observe patient for 45-60 min. If no adverse events, administer second dose</td>
</tr>
<tr>
<td><strong>COWS &gt;13</strong></td>
<td></td>
</tr>
<tr>
<td>Buprenorphine-naloxone (buprenorphine [dosed with naloxone])</td>
<td>• 8 mg, Tab-SL, sublingual, once, now 0.5 mg naloxone per each 2 mg buprenorphine</td>
</tr>
<tr>
<td><strong>Discharge prescriptions</strong></td>
<td></td>
</tr>
<tr>
<td>Buprenorphine-naloxone (buprenorphine-naloxone 8 mg-2 mg sublingual tablet)</td>
<td>• =1 tablet, sublingual, BID&lt;sup&gt;e&lt;/sup&gt;, #14 tablets, refill 0</td>
</tr>
<tr>
<td>Naloxone intranasal (take-home supply)</td>
<td></td>
</tr>
<tr>
<td>Buprenorphine-naloxone (buprenorphine-naloxone 8 mg-2 mg sublingual tablet)</td>
<td>• =1 tablet, sublingual, BID, #20 tablets, refill 0</td>
</tr>
<tr>
<td><strong>Laboratory</strong></td>
<td></td>
</tr>
<tr>
<td>Drugs of abuse profile (urine drug screen)</td>
<td>• Urine</td>
</tr>
<tr>
<td>Comprehensive metabolic panel</td>
<td>• Blood</td>
</tr>
<tr>
<td>Order urine pregnancy test for females aged 15-55 years</td>
<td>• Urine</td>
</tr>
<tr>
<td><strong>Consults</strong></td>
<td></td>
</tr>
<tr>
<td>Consult to social services</td>
<td>• Other (use special instructions); evaluate opioid use disorder, treatment, and referral in the emergency department</td>
</tr>
</tbody>
</table>

<sup>a</sup>UED: university emergency department.  
<sup>b</sup>COWS: Clinical Opioid Withdrawal Scale.  
<sup>c</sup>Tab-SL: tablet-sublingual.  
<sup>d</sup>PRN: pro re nata (as needed).  
<sup>e</sup>BID: bis in die (twice daily).  

Concomitantly, prompted by a nursing order in the OUD order set, the nursing staff will assess the presence and severity of opioid withdrawal symptoms using the Clinical Opioid Withdrawal Scale (COWS) [20]. If the COWS score is between 8 and 13 (moderate withdrawal), a 4 mg-1 mg sublingual dose of buprenorphine-naloxone will be offered to the patient for induction. The patient will be observed for 1 hour, and a second 4 mg/1 mg dose will be provided if needed based on symptomatic response and repeated COWS score assessment. If the COWS score is >13 (severe withdrawal), an 8 mg-2 mg sublingual dose of buprenorphine-naloxone will be offered for induction, and the patient will be observed for symptomatic improvement [21]. Buprenorphine-naloxone administered in the ED will be stored in Omnicell automated medication-dispensing cabinets located in the ED for efficient access and delivery to patients in withdrawal. If the patient is agreeable to engaging in outpatient addiction treatment, a 10-day prescription for buprenorphine-naloxone, 8 mg-2 mg sublingual BID (twice daily), will be offered for take-home induction bridge to outpatient MOUD (10 days has been determined to be within the current average time to follow-up for local MOUD clinics). In addition, a take-home intranasal naloxone kit will be provided along with instructions on use and general overdose education (Multimedia Appendix 1). The buprenorphine-naloxone prescription will be filled on site by the hospital pharmacy, allowing the patient to be discharged with the medication in hand. The take-home naloxone kit and supply of buprenorphine-naloxone will also be provided at no cost to enrolled patients. The study cost of buprenorphine-naloxone is US $6 per pill. The intranasal
naloxone kits are provided free of charge to our department by the Jefferson County Health Department.

**Medical Clearance**

Before buprenorphine-naloxone induction and activation of bedside referral services, patients will undergo routine medical screening and clearance consisting of a comprehensive metabolic panel, urine drug test (UDT), and a urine pregnancy test in female patients with childbearing potential. Although the only true contraindication to buprenorphine is a hypersensitivity reaction to the medication, dose modification in the setting of hepatic dysfunction will be considered. Urine drug screening will allow clinicians to detect patients who are malingering and/or seeking to divert buprenorphine-naloxone. Pregnant patients with OUD in active withdrawal or seeking addiction care receive a consultation for psychiatric addiction services in the ED. Preexisting institution-specific services offer focused monitoring and intervention for this specific population to minimize the risk of neonatal abstinence syndrome and miscarriage and optimize pregnancy outcomes [22]. As such, pregnant patients will be excluded from this study.

**Plans to Mitigate the Risk of Diversion**

Only buprenorphine-naloxone, which has a lower risk of diversion than methadone, will be administered by ED staff per DATA 2000 waiver training regulations [23]. The Alabama PDMP will be checked before each buprenorphine-naloxone prescription (as described above). Each prescription’s amount and duration will be documented in the EHR and communicated to the RRC and subsequently to the receiving Alabama Department of Mental Health (ADMH) certified treatment facility. Physician engagement and education (see Physician Education and Engagement section) include standard DATA 2000 waiver training, which encompasses UDT interpretation. Therefore, the UDT collection will be included in the OUD EHR order set, and result consideration will be advised per departmental protocol before administering or prescribing buprenorphine-naloxone. Likewise, consideration of the PDMP will be strongly encouraged for any provider prescribing buprenorphine-naloxone. No specific restrictions or guidelines will be provided during this study concerning either the PDMP or UDT analysis beyond the standard DATA 2000 training. Rather, they are additional resources available to our physicians for both OUD diagnosis and evaluation of suspected diversion.

ADMH-certified facilities that receive OUD patients are required to independently develop, maintain, and implement diversion control plans. We will also limit the duration of the buprenorphine-naloxone prescription to 10 days, with an option for a subsequent as needed 7-day refill to balance 2 potential factors inherent to local substance abuse treatment referral: (1) waiting time before outpatient addiction services are available and (2) the risk of diversion associated with prescriptions of prolonged duration.

Despite linkage support and efforts, there may be occasions when a patient fails to link to outpatient MOUD. In these circumstances, the patient may return to the ED seeking reengagement in the program and a MOUD refill. In these instances, research staff engagement can be used to identify and potentially overcome specific barriers to the initial treatment plan. Recidivism will also be monitored. However, the decision to provide a return patient with a refill prescription will be determined on an individual basis and at the discretion of the prescribing physician.

**Physician Education and Engagement**

In the state of Alabama, 2 hours of opioid-specific continuing medical education are required every 2 years, which aligns with the goal of UAB-attending physicians to obtain DATA 2000 waiver training. In the months leading up to the project start date, the principal investigator provided brief presentations at monthly emergency medicine faculty meetings to describe the evidence supporting MOUD initiation in the ED—the high incidence of opioid overdose and death in Jefferson County—and advocate for the importance of a concerted initiative centered around the needs of patients with OUD. In addition, coauthors (LAW and JJH) developed 4 hours of didactic material to deliver to emergency medicine residents in a single, half-day format, including presentations, small group case discussions, simulations, and a presentation by leadership from the RRC. In an effort to encourage learner engagement and provide a general overview and introduction of the in-person didactic, we created and released a peer-reviewed podcast that is available on iTunes and Soundcloud [24,25]. Finally, as a public health intervention consistent with the mission and values of the UAB Department of Emergency Medicine, the Emergency Medicine Department Chair mandated the faculty completion of X-waiver training.

**Community Partnerships**

The RRC is a local addiction coordinating center that facilitates intake assessment and substance abuse referral and placement for all-comers, regardless of insurance status. Following patient intake and assessment, the RRC uses local resources and networks to direct the patient to the most appropriate definitive MOUD continuing care option available. The engagement of the RRC as a local clinical referral partner in this project was a crucial initial step. Assisting individuals with substance abuse in navigating the treatment system is an RRC mission. Formalizing a referral process with the RRC as a part of this project fulfilled a critical need for our patients and helped them fulfill their vision as a valuable community resource. As described above, when an OUD patient is identified by ED staff, the RRC will be engaged by research staff to dispatch a trained peer recovery specialist to ensure a warm handoff between ED and RRC team members and initiate referral procedures for follow-up. A warm handoff is a direct transfer of care between 2 members of a health care team that typically occurs in front of the patient. This seamless transition in care has been shown to decrease the time to enrollment in recovery programs and improve recovery prospects [26]. Following study enrollment and contact by research staff, the peer recovery specialists meet directly with the patient in the ED, creating a direct personal link to the RRC for subsequent linkage to care steps. Peer recovery specialists or peer navigators are individuals in sustained recovery from OUD who have undergone a certification process involving 40 hours of specialized substance abuse training. Peer navigators are funded by the RRC, which...
is supported by the local county health department and the Birmingham Crisis Center. The peer navigators provide direct assistance to the patients to facilitate access and intake at the RRC (eg, physically accompanying them on the same day to RRC intake or sharing contact information to use next-day contact to ensure RRC follow-up). RRC staff will perform a detailed intake assessment, including the American Society for Addiction Medicine and Government Performance and Results Modernization Act (GPRA) assessments, and initiate a referral to an ADMH-certified treatment facility for the continuation of MOUD. ADMH-certified treatment facilities deliver comprehensive psychosocial services, including drug counseling, recovery support services, and behavioral therapies, in addition to medication. Within our local community, we currently have 4 inpatient facilities that provide opioid detoxification and rehabilitation services, including MOUD, comprising approximately 200-plus inpatient beds with varying wait times for intake (immediate to 1 week). In addition, we have 5 outpatient facilities, several of which overlap with the aforementioned inpatient facilities, with additional capacity for weekly (during new induction) and monthly (maintenance) MOUD services. Wait times for outpatient linkage averages 3 days to 2 weeks, depending on the facility.

In addition, per RRC standard practice, peer navigators will continue to facilitate the patient’s care until the patient has obtained MOUD clinic follow-up or the patient no longer wishes to be contacted by the peer navigator or is otherwise unreachable.

Letters of commitment from referral partners (see the Data Collection and Evaluation section) allowed communication between research staff, the RRC, and MOUD referral partners to be bidirectional. In addition, the UAB research faculty participate in the RRC oversight committee, which regularly engages community substance abuse treatment partners in the discussion and monitoring of local MOUD resources.

Data Collection and Evaluation

The GPRA of 1993, which was revised in 2010, requires SAMHSA grantees to collect and report performance data using standardized measurement tools [27]. GPRA tools gather extensive patient-specific information, including basic demographics, substance use and abuse, mental health and physical health functioning, and other key variables. The baseline GPRA data collection will be coordinated and completed by the dedicated research staff during the initial ED visit. In-person follow-up data collection interviews at 3 and 6 months postintake and at discharge from addiction treatment will be collected by a contracted third-party vendor, Community Tracking Services (CTS). Follow-up data (eg, current address and contact information) will be obtained through data-sharing follow-up efforts led by referral facilities, including the RRC. Discharge from addiction treatment as an event is technically determined by each respective MOUD clinic and will align with individual patient treatment plans. However, for this project’s purposes, it will be defined as when the patient is no longer maintained on MOUD maintenance therapy, including concomitant psychosocial support services, and therefore no longer an active patient at the MOUD clinic.

Contact barriers for obtaining follow-up data have been considered. Research and CTS staff will rely heavily on MOUD support partners and referral centers, namely, the RRC and the RRC peer navigators, to ensure accurate and up-to-date contact information for enrolled patients is maintained. In addition, contact information for family and friends will also be obtained with patient consent at the time of enrollment. MOUD clinic referral partners will also collect additional contact information.

Project data will be stored in a secure REDCap (Research Electronic Data Capture) database housed on UAB-encrypted servers [28]. Per requirements, project data are also entered and stored in SAMHSA’s Performance Accountability and Reporting System, a web-based data entry, reporting, and technical assistance portal. Formal institutional review board (IRB) protocols and data use agreements have been executed between UAB and non-UAB partners to ensure confidentiality, human subject protection, and Health Insurance Portability and Accountability Act compliance. Written consent will be obtained according to federal confidentiality law and regulation 42 Code of Federal Regulations, Part 2, to exchange health information between UAB, the RRC, and ADMH-certified MOUD treatment facilities to coordinate care delivery. Program evaluation activities will measure (1) the fidelity of program delivery to the timeline, (2) delivery of DATA 2000 training as targeted, (3) data demonstrating implementation of evidence-based practices (eg, MOUD and peer support services), and (4) GPRA data collection. MOUD referral partners and CTS are permitted to collect and share data under the provision of a signed letter of commitment, which describes the respective data collection and sharing roles and responsibilities. Letters of commitment were included in the IRB protocol.

The outcomes and measurable project goals are listed in Table 1. Progress toward the first goal will be evaluated by comparing the percentage of documented DATA 2000 waivers with the census of ED physicians with current DEA certificates. Achievement of goal 2 will be assessed by electronically tracking when the protocols are used and the EHR review. Goal 3 will be evaluated based on GPRA interview data collected at the time of referral to the RRC. Goals 4 and 5 will be evaluated by comparing 3-month, 6-month, and discharge follow-up information from year to year. Goal 5 will also be evaluated by comparing opioid-related deaths at the end of years 2 and 3 from Jefferson County Coroner reports, with the average between 2014 and 2017 baseline years.

Data Analysis and Management

Analysis of outcome data will include descriptive statistics, including tables that summarize quantitative data and the number of clients who complete treatment compared with those who do not. We will test for differences in a variety of demographic and socioeconomic variables, including racial and ethnic groups. Client data will be analyzed at the time of enrollment, 3 and 6 months post enrollment, and at the time of discharge from local substance abuse treatment facilities. Program outcomes will be monitored quarterly. Qualitative data will be analyzed according to the project evaluator’s procedures, as appropriate for the variable collected.

https://www.researchprotocols.org/2021/4/e18734

JMIR Res Protoc 2021 | vol. 10 | iss. 4 | e18734 | p.429 (page number not for citation purposes)
Results

Funded in February 2019, this study is in the active enrollment and data collection phase. The UAB IRB approved it in June 2019, and data collection and enrollment began on July 7, 2019. The project end date is projected to be February 26, 2022. As of February 2019, we have enrolled 79 participants. We anticipate the results of the public health intervention to be published in the summer of 2022.

Discussion

Overview

Across the United States, the opioid epidemic continues to rage, disproportionately affecting particular areas of the nation, including Appalachia. Appalachia is characterized by a high poverty rate (16.3% as compared with 14.6% nationally) and limited access to subspecialty care (42% of the Appalachian population is rural compared with 20% of the national population), and Appalachian counties have more health care costs coupled with coverage and access disparities than the rest of the United States [29]. Alabama is a non-Medicaid expansion state, which further exacerbates these health care disparities [30].

These characteristics highlight the potential impact of an ED-based intervention for patients with OUD at UAB, the largest and only tertiary care academic medical center in the state of Alabama. They also highlight the necessity of collaboration and community partnerships for a public health intervention such as this to be successful. Many individuals who require addiction care comprise an underserved population with limited access to health care. Our local community partner, the RRC, coordinates care for over 1000 patients with OUD or SUD annually and refers their clients for MOUD, demonstrating an efficient model of care. The linkage of this community resource directly to OUD patients in the ED is critical to ensure continuity of care. EDs have previously served as effective venues to raise awareness of and linkage to care for similar stigmatizing and deadly conditions—specifically HIV and hepatitis C virus [31,32]. The successful implementation of the public health intervention described herein further demonstrates that the ED setting represents a unique opportunity to engage with the community and improve the health of the public [33].

Future Directions

As we implement the ED-MOUD protocol, a greater number of individuals with OUD will be identified and referred to treatment, increasing the demand for MOUD. As there is a shortage of addiction providers in our region, there is an urgent need to increase the number of clinicians with DATA 2000 training who can effectively provide MAT longitudinally in collaboration with UAB addiction specialists.

An integrated hub and spoke opioid treatment network was developed in Vermont, which may be appropriate for modification and application in our region [34]. In the Vermont model, hub clinics with addiction medicine expertise assess patients’ medical and psychiatric needs at intake, induce patients with buprenorphine, and determine the most appropriate treatment placement. Once patients are stable on buprenorphine, they are referred to a spoke provider. Spokes have direct access to hubs for consultation, screening, and MOUD delivery, and transfers between spokes and hubs can be bidirectional. Any waivered physician can become a spoke provider, hoping that with the extra support of the hub, the MOUD capacity will increase. In Alabama, a non-Medicaid expansion state, we will need to modify the Vermont spoke and hub model to be financially sustainable. In Jefferson County, Federally Qualified Health Centers (FQHCs), community-based health care providers that receive funds from the Health Resources and Services Administration Health Center program, provide primary care services in underserved areas and have the potential to act as spokes in a modified Vermont model [35].

A recent additional grant from SAMHSA will enable the investigative team to establish a Provider Clinical Support System Data 2000 waiver training program at our institution along with a program to train primary care clinicians to deliver MOUD to patients with OUD. This program will facilitate collaboration with FQHCs in Jefferson County to provide addiction training, oversight, and consultation to expand primary care providers’ scope of practice to include MOUD. In addition, we are developing a telemedicine-assisted consultation model to support primary care providers as they deliver care to individual patients. We hope that these additional training programs and models will expand MOUD capacity in Alabama, similar to the Vermont model.

We are also exploring opening a MOUD bridge clinic at UAB to manage patients between the time of identification in the ED and linkage to longitudinal addiction treatment. Such a clinic would also provide an opportunity to incorporate multidisciplinary engagement, including medical social workers and case management services, strategically identifying and addressing patients’ social determinants of health to facilitate treatment of OUD. It will also provide a supervised venue for training physicians in the complexities of MOUD, equipping them to continue these addiction practices as they are subsequently hired throughout the state in rural communities.

Conclusions

Traditionally a niche reserved for psychiatrists or addiction medicine specialists, treatment of OUD with MOUD is expanding to emergency clinicians, primary care clinicians, and other health care providers. EDs are the tip of the spear where societal problems meet health care, and the intersection between the opioid epidemic and the emergency care system highlights this reality. Initiating MOUD in the ED is an evidence-based approach for improving health outcomes in OUD. Community collaboration is critical for developing a feasible, cost-effective, and sustainable approach to combat the opioid epidemic in resource-limited regions.
Acknowledgments

SAMHSA provided funding for this project. The funder predefined a standardized data collection instrument as a component of the data collected according to the GPRA, but otherwise, the funder had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

Authors’ Contributions

LAW is a core research and educational faculty member and a major contributor in writing the manuscript and cocreating the emergency medicine resident OUD education module. LL is an addiction medicine specialist and ED liaison for MOUD linkage support services; she also contributed to the manuscript’s editing. JBR is an emergency research staff member and is integral in grant submission and current program and data maintenance. He also contributed to the editing of the manuscript. JJJ created and codirected the emergency medicine resident OUD education module and contributed to the manuscript’s editing. RMS is the core research faculty member responsible for logistical program implementation, and she contributed to the writing and editing of the manuscript. MCD is a core educational faculty member and created the podcast for the emergency medicine resident OUD educational module. JB provided patient demographic data and assisted in the creation of the custom ED EHR opioid order set. EPH is the project’s principal investigator and was a major contributor in writing the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Emergency departments’ medications for opioid use disorder induction flowsheet.

[DOCX File, 49 KB - resprot_v10i4e18734_app1.docx ]

References

15. IC-D10-CM Codes.10. URL: [accessed 2021-04-14]
24. 5 x 5: evidence based medicine. Apple Podcasts Preview. URL: [accessed 2020-10-23]
25. "We didn't start the fire": opioid use disorder in the emergency department. 5 X 5: Evidence Based Medicine. URL: [accessed 2021-04-14]
27. GPRA Modernization Act of 2010. Substance Abuse Mental Health Services Administration. URL: [accessed 2020-10-11]
35. Federally qualified health centers. Health Resources and Services Administration. 2018. URL: [accessed 2019-10-10]

Abbreviations

ADMH: Alabama Department of Mental Health
BNI: brief negotiation interview
COWS: Clinical Opioid Withdrawal Scale
CTS: Community Tracking Services
DATA: Drug Addiction Treatment Act
DEA: Drug Enforcement Administration
DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)
ED: emergency department
EHR: electronic health record
FQHC: Federally Qualified Health Centers
GPRA: Government Performance and Results Modernization Act
IRB: institutional review board
MAT: medication-assisted treatment
MOUD: medications for opioid use disorder
OUD: opioid use disorder
PDMP: prescription drug monitoring program
RRC: Recovery Resource Center
SAMHSA: Substance Abuse and Mental Health Services Administration
SUD: substance use disorder
UAB: University of Alabama at Birmingham
UDT: urine drug test

Edited by G Eysenbach; submitted 15.03.20; peer-reviewed by J Hoppe, J Mihanovic, C Freiermuth; comments to author 21.08.20; revised version received 07.10.20; accepted 03.11.20; published 29.04.21.

Please cite as:
Development of an Emergency Department–Based Intervention to Expand Access to Medications for Opioid Use Disorder in a Medicaid Nonexpansion Setting: Protocol for Engagement and Community Collaboration
JMIR Res Protoc 2021;10(4):e18734
URL: https://www.researchprotocols.org/2021/4/e18734
doi:10.2196/18734
PMID:33913818

©Lauren A Walter, Li Li, Joel B Rodgers, Jennifer J Hess, Rachel M Skains, Matthew C Delaney, James Booth, Erik P Hess. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.