# Contents

## Protocols

Impact of a Mobile App (LoAD Calc) on the Calculation of Maximum Safe Doses of Local Anesthetics: Protocol for a Randomized Controlled Trial *(e53679)*

Pietro Fubini, Georges Savoldelli, Tal Beckmann, Caroline Samer, Mélanie Suppan. ................................................................. 4

Professional Development to Improve Responsible Beverage Service Training: Formative Research Results and Protocol for a Randomized Controlled Trial *(e49680)*

W Woodall, David Buller, Robert Saltz, Lila Martinez. ................................................................. 14

Novel Strategy to Assess the Neurotoxicity of Organic Solvents Such as Glycol Ethers: Protocol for Combining In Vitro and In Silico Methods With Human-Controlled Exposure Experiments *(e50300)*

Nancy Hopf, Laura Suter-Dick, Jörg Huwyler, Myriam Borgatta, Lucie Hegg, David Parnies, Hélène Paschoud, Ramya Puligilla, Elena Reale, Sophie Werner, Marie-Gabrielle Zurich. ................................................................. 30

Implementation of a Primary Prevention Program for Posttraumatic Stress Disorder in a Cohort of Professional Soldiers (PREPAR): Protocol for a Randomized Controlled Trial *(e47175)*

Emmanuelle Le Barbenchon, Marion Trousselard, Sonia Pellissier, Mathilde Moisson-Baudé, Philippe Chachignon, Pierre Bouny, Emma Touré Cuq, Sandrine Jacob, Cécile Vigier, Maud Hidalgo, Damien Claverie, Anais Duffaud. ................................................................. 47

Investigation of the Association Between e-Cigarette Smoking and Oral Mucosal Health Status Among Young People: Protocol for a Case-Control Trial *(e53644)*

Siyuan Cheng. ................................................................. 59

Leveraging Community Health Workers and a Responsive Digital Health System to Improve Vaccination Coverage and Timeliness in Resource-Limited Settings: Protocol for a Cluster Randomized Type 1 Effectiveness-Implementation Hybrid Study *(e52523)*

Lavanya Vasudevan, Jan Ostermann, Nathan Thielman, Joy Baumgartner, David Solomon, Anna Mosses, Amy Hobbie, Nicole Hair, Chen Liang, Marco van Zwetselaar, Sayoki Mfinanga, Esther Ngadaya. ................................................................. 65

Testing an Evidence-Based Self-Help Program for Infertility-Related Distress: Protocol for a Randomized Controlled Trial *(e52662)*

Jennifer Gordon, Megan Poulter, Ashley Balsom, Tavis Campbell. ................................................................. 79

Effects of a Powered Ankle-Foot Prosthesis and Physical Therapy on Function for Individuals With Transfemoral Limb Loss: Rationale, Design, and Protocol for a Multisite Clinical Trial *(e53412)*

Jason Maikos, Alison Puziner, Brad Hendershot, David Herilhy, John Chomack, Michael Hyre, Samuel Phillips, Alexis Sidiropoulos, Christopher Dearth, Leif Nelson. ................................................................. 90
Exploring Medical Career Choice to Better Inform Swiss Physician Workforce Planning: Protocol for a National Cohort Study (e53138)
Milena Abbiati, Mathieu Nendaz, Bernard Cerutti, Monika Brodmann Mäder, Giatgen Spinas, David Vicente Alvarez, Douglas Teodoro, Georges Savoldelli, Nadia Bajwa. .................................................................................................................................................. 109

Participatory Development and Assessment of Audio-Delivered Interventions and Written Material and Their Impact on the Perception, Knowledge, and Attitudes Toward Leprosy in Nigeria: Protocol for a Cluster Randomized Controlled Trial (e53130)
Ngozi Murphy-Okpala, Tahir Dahiru, Anna van’t Noordende, Carolin Gunesch, Joseph Chukwu, Charles Nwafor, Suleiman Abdullahi, Chukwuma Anyaile, Ugochinyere Okereke, Anthony Meka, Chinwe Eze, Okechukwu Ezeakile, Ngozi Ekeke. .................................................................................................................................................. 119

Adaptive Intervention to Prevent Respiratory Illness in Cerebral Palsy: Protocol for a Feasibility Pilot Randomized Controlled Trial (e49705)
Alyssa Fleischman, Carlos Lerner, Heidi Kloster, Paul Chung, Thomas Kitzner, Christopher Cushing, Danielle Gerber, Barbara Katz, Gemma Warner, Kristina Singh-Verdeflor, Roxana Delgado-Martinez, Lorena Porras-Javier, Siem Ia, Teresa Wagner, Mary Ehlenbach, Ryan Coller. .................................................................................................................................................. 130

Efficacy and Safety of the Natural Killer T Cell–Stimulatory Glycolipid OCH-NCNP1 for Patients With Relapsing Multiple Sclerosis: Protocol for a Randomized Placebo-Controlled Clinical Trial (e46709)
Tomoko Okamoto, Takami Ishizuka, Reiko Shimizu, Yasuko Asahina, Harumasa Nakamura, Yoko Shimizu, Yoichiro Nishida, Takanori Yokota, Youwei Lin, Wako Sato, Takashi Yamamura. .................................................................................................................................................. 145

Transcranial Magnetic Stimulation of the Default Mode Network to Improve Sleep in Individuals With Insomnia Symptoms: Protocol for a Double-Blind Randomized Controlled Trial (e51212)
Lindsey Hildebrand, Alisa Huskey, Natalie Dailey, Samantha Jankowski, Kymberly Henderson-Arredondo, Christopher Trapani, Salma Patel, Allison Chen, Ying-Hui Chou, William Killgore. .................................................................................................................................................. 154

Epidemiology of Syphilis in Pregnancy and Congenital Syphilis in Brazil and the Risk or Associated Factors: Protocol for a Systematic Review (e50702)
Yago Pinheiro, Jamilli Dantas, Jose Holanda, Anilma Feitosa, Richardson Augusto Rosendo da Silva. .................................................................................................................................................. 167

Reward Learning as a Potential Mechanism for Improvement in Schizophrenia Spectrum Disorders Following Cognitive Remediation: Protocol for a Clinical, Nonrandomized, Pre-Post Pilot Study (e52505)
Frances Dark, Graham Galloway, Marcus Gray, Matteo Celli, Veronica De Monte, Victoria Gore-Jones, Gabrielle Ritchie. .................................................................................................................................................. 174

Education of Patients With Atrial Fibrillation and Evaluation of the Efficacy of a Mobile Virtual Patient Environment: Protocol for a Multicenter Pseudorandomized Controlled Trial (e45946)
Panagiotis Antoniou, Eleni Dafli, George Giannakoulas, Gaukhar Igimbayeva, Olga Visternichan, Serhii Kyselov, Ivetta Lykhasenko, Dmytro Lashkul, Ilia Nadareishvili, Sero Tabagari, Panagiotis Baidis. .................................................................................................................................................. 186

Development and Testing of an Electronic Diabetes Diary Integrated With a Hospital Information System for Individuals With Type 2 Diabetes Mellitus: Protocol for a Mixed Methods Study (e50732)
Geena Skaria, Bhageerathy Reshmi, Sabu K M, Sahana Shetty, Vani Lakshmi R. .................................................................................................................................................. 198

The Effects of the Processing of Positive Memories Technique on Posttrauma Affect and Cognitions Among Survivors of Trauma: Protocol for a Daily Diary Study (e51838)
Talya Greene, Ateka Contractor, Sheila Dicker-Oren, Andrea Fentem, Sharon Szinman. .................................................................................................................................................. 207

Microfragmented Fat and Biphasic Calcium Phosphates for Alveolar Cleft Repair: Protocol for a Prospective, Nonblinded, First-in-Human Clinical Study (e42371)
Diandra Natsir Kalla, Salem Alkaabi, Abul Faizi, Andi Tajrin, Rifaat Nurrahma, Werner Müller, Heinz Schröder, Xiaohong Wang, Tymour Forouzanfar, Marco Held, Muhammad Ruslin. .................................................................................................................................................. 218
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue and Mental Illness Symptoms in Long COVID: Protocol for a Prospective Cohort Multicenter Observational Study (e51820)</td>
<td>Ligia Pires, Cláudia Reis, Ana Mesquita Facção, Armin Moniri, Ana Marreiros, Marta Drummond, Joana Berger-Estilita.</td>
<td>227</td>
</tr>
<tr>
<td>The Impact of Forced Separations Between Women and Their Pets in Domestic Violence Situations and the Effectiveness of Crisis Response: Protocol for a Conceptual Framework (e52067)</td>
<td>Jasmine Montgomery, Janice Lloyd, Zhanming Liang.</td>
<td>238</td>
</tr>
<tr>
<td>Mapping Respiratory Health Digital Interventions in South and Southeast Asia: Protocol for a Scoping Review (e52517)</td>
<td>Laura Evans, Jay Evans, Monica Fletcher, Adina Abdullah, Zakuddin Ahmed.</td>
<td>253</td>
</tr>
<tr>
<td>Seasonal Malaria Chemoprevention Therapy in Children Up To 9 Years of Age: Protocol for a Cluster-Randomized Trial Study (e51660)</td>
<td>Mahamoudou Toure, Jeffrey Shaffer, Daouda Sanogo, Soumba Keita, Moussa Keita, Fousseyni Kane, Bourama Traore, Djeneba Dabitao, Aissata Kone, Cheick Doumbia, Joseph Keating, Joshua Yukoch, Helle Hansson, Alyssa Barry, Mahamadou Diakité, Michael Alifrangis, Seydou Doumbia.</td>
<td>281</td>
</tr>
</tbody>
</table>

**Corrigenda and Addendas**

**Correction: Mobile Phone Technology for Preventing HIV and Related Youth Health Problems, Sexual Health, Mental Health, and Substance Use Problems in Southwest Uganda (Youth Health SMS)- Protocol for a Pilot Randomized Controlled Trial (e55725)**
Philip Kreniske, Olive Namuyaba, Robert Kasumba, Phionah Namatovu, Fred Ssewamala, Gina Wingood, Ying Wei, Michele Ybarra, Charlotte Oloya, Costella Tindyebwa, Christina Ntulo, Vincent Mjujane, Larry Chang, Claude Mellins, John Santelli. | 247  |

**Correction: A Machine Learning Model to Predict Patients’ Adherence Behavior and a Decision Support System for Patients With Metastatic Breast Cancer: Protocol for a Randomized Controlled Trial (e55928)**
Marianna Masiero, Gea Spada, Virginia Sanchini, Elisabetta Munzone, Ricardo Pietrobon, Lucas Teixeira, Mirtha Valencia, Aline Machiavelli, Elisa Fragale, Massimo Pezzolato, Gabriella Pravettoni. | 249  |

**Correction: The Development of a UK Culturally Adapted and Modified Version of the Person Attuned Musical Interactions Manual: Protocol for a 2-Phase Mixed Methods Study (e56085)**
Bryony Waters, Martin Orrell, Orli McDermott. | 251  |
Impact of a Mobile App (LoAD Calc) on the Calculation of Maximum Safe Doses of Local Anesthetics: Protocol for a Randomized Controlled Trial

Pietro Elias Fubini1, MD; Georges Louis Savoldelli1, MEd, MD; Tal Sara Beckmann1, MD; Caroline Flora Samer2, MAS, MD; Mélanie Suppan1, MSc, MD

1Division of Anesthesiology, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Geneva University Hospitals and Faculty of Medicine, Geneva, Switzerland
2Clinical Pharmacology and Toxicology Division, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Geneva University Hospitals and Faculty of Medicine, Geneva, Switzerland

Corresponding Author:
Mélanie Suppan, MSc, MD
Division of Anesthesiology
Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine
Geneva University Hospitals and Faculty of Medicine
Rue Gabrielle-Perret-Gentil 4
Geneva, 1211
Switzerland
Phone: 41 223723311
Email: melanie.suppan@hcuge.ch

Abstract

Background: Local anesthetics (LAs) are regularly used to alleviate pain during medical or surgical procedures. Their use is generally considered safe, but exceeding the maximum recommended doses can lead to LA systemic toxicity, a rare but potentially lethal complication. Determining maximum safe doses is therefore mandatory before performing local anesthesia, but rules are often unclear and the factors affecting dose calculation are numerous. Mobile health apps have been shown to help clinical decision-making, but most currently available apps present significant limitations. The Local Anesthetics Dose Calculator (LoAD Calc) app was designed to overcome these limitations by taking all relevant parameters into account. Before deploying this app in a clinical setting, it should be tested to determine its effectiveness and whether clinicians would be willing to use it.

Objective: The primary objective will be to evaluate the effectiveness of the LoAD Calc app through written simulated cases. The secondary objective will be to determine whether physicians find this app easier, faster, and safer than the methods they generally use.

Methods: We describe a parallel-group randomized controlled trial protocol. Anesthesiologists working at the Geneva University Hospitals will be invited to participate. Participants will be asked to compute the maximum dose of LA in 10 simulated clinical cases using 3 different LAs. The maximum safe dose will be determined manually using the same calculation rules that were used to develop LoAD Calc, without using the app itself. An overdose will be considered any dose higher than the correct dose, rounded to the superior integer, while an underdose will be defined as the optimal calculated dose minus 20%, rounded to the inferior integer. Randomization will be stratified according to current position (resident vs registrar). The participants allocated to the LoAD Calc (experimental) group will use the LoAD Calc app to compute the maximum safe LA doses. Those allocated to the control group will be asked to use the method they generally use. The primary outcome will be the overall overdose rate. Secondary outcomes will include the overdose rate according to ideal and actual body weight and to each specific LA, the overall underdose rate, and the time taken to complete these calculations. The app’s usability will also be assessed.

Results: A sample size of 46 participants will be needed to detect a difference of 10% with a power of 90%. Thus, a target of 50 participants was set to allow for attrition and exclusion criteria. We expect recruitment to begin during the winter of 2023, data analysis in the spring of 2024, and results by the end of 2024.

Conclusions: This study should determine whether LoAD Calc, a mobile health app designed to compute maximum safe LA doses, is safer and more efficient than traditional LA calculation methods.

International Registered Report Identifier (IRRID): PRR1-10.2196/53679
Introduction

Background

Local anesthetics (LAs) are used daily by physicians to perform minor procedures. While the doses they use are generally limited, anesthesiologists often use higher doses to perform regional anesthesia techniques [1]. While the advantages of such techniques are undeniable, using high LA doses increases the risk of local anesthetic systemic toxicity (LAST), a potentially lethal complication associated with the use of these agents [2]. The actual incidence of LAST is unknown since most minor symptoms are not specific and because LAST awareness varies considerably between practitioners [3,4]. The incidence reported in scientific studies varies from 0.04 to 1.8 per 1000 regional anesthesia procedures but is probably underestimated [5,6]. The main risk factors seem to be inadvertent intravascular injections and inappropriately large doses [7,8].

Prevention of intravascular injection can be achieved by ultrasound guidance and careful aspiration during the procedure, while adequate calculation of the maximum dose of LA before administration is the best way to avoid incorrect doses.

Although different guidelines have been created to help clinicians calculate the maximum safe LA doses, quickly and reliably determining such doses often proves difficult in clinical practice [9]. Many anesthesiologists rely on mental calculation (with or without a pen and paper aid), and some use calculators. These methods are, however, often challenging and inaccurate, especially if LA mixtures are used or when patients present multiple comorbidities [10]. More advanced solutions have been developed to support LA dosage calculation, such as the nomogram created by Williams and Walker [11]. The main limitation of this solution is that the nomogram must always be at hand. Moreover, specificities such as ideal body weight (IBW) calculation and adaptation in the case of relevant comorbidities are indicated but not directly integrated into dose determination.

To facilitate the calculation of safe maximum LA doses, a mobile health (mHealth) app, Local Anesthetics Dose Calculator (LoAD Calc), was developed at the Geneva University Hospitals [12]. This app takes all relevant parameters (IBW and actual weight, height, age, medications, and comorbidities) into account and allows the use of a mixture of 2 different LAs. Since smartphones have widely replaced older paging systems and are therefore always at hand, this mHealth app could be an appropriate solution to enable anesthesiologists to efficiently compute safe maximum LA doses.

Objectives

This study protocol follows the hypothesis that LoAD Calc, an mHealth app designed to help clinicians calculate maximum safe LA doses, is safer and more effective than traditional methods. Thus, the primary objective will be to evaluate the effectiveness of the LoAD Calc app by using it to compute the maximum single doses of LA in written simulated cases. The secondary objective will be to determine whether physicians find this app easier, faster, and safer than the methods they generally use.

Methods

Ethical Considerations

A synopsis of the study protocol was presented to the regional ethics committee (Commission Cantonale d’Ethique de la Recherche [CCER]). This committee confirmed that this project does not fall within the scope of the Swiss Federal Act on Research involving Human Beings and issued a “declaration of no objection” (CCER 2022-01577) [13]. This study protocol does not fall within the scope of the Swiss Federal Act on Research involving Human Beings [13]. It will nevertheless be presented to the regional ethics committee to ascertain that no important or relevant ethical consideration was omitted.

Participants will be told that participation is entirely voluntary, that there will be no consequence if they refuse to participate, and that they will be able to withdraw at any time without explanation. All participants will be asked to sign an electronic consent form immediately after logging in. There is no compensation for participation in the study.

Study Design

This will be a monocentric, parallel-group, randomized controlled trial based on clinical vignettes. The protocol was developed according to the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement (Multimedia Appendix 1) [14]. Given its design, the investigators will not be blinded as to the intervention. Nevertheless, participants will not be informed that there are 2 different arms and will not be told the exact outcomes studied, even though they will be provided with general information regarding the study. In addition, the data analyst will be blinded as to participant allocation by renaming the groups before sending data for statistical analysis. Randomization will be stratified according to current position (resident vs registrar).

Results will be reported according to the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and onLine TeleHealth) guidelines [15]. Relevant elements of the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) will be included since web-based questionnaires will be used in the course of this study [16].

Clinical Vignettes

A total of 10 clinical vignettes will be developed for the purpose of this study. These vignettes will describe clinical cases
requiring the use of LAs for regional anesthesia. We will include 3 of the most commonly used LAs in these vignettes: lidocaine, levobupivacaine, and ropivacaine. Some vignettes will ask the participant to use LA mixtures, and several will include comorbidities or medications requiring dose adaptations.

For each vignette, 3 authors will be required to determine the maximum dose of LA the simulated patient should receive according to the rules used to develop the LoAD Calc app [12], without using the app. These rules, which are derived from the scientific literature, are summarized in Textbox 1. They were reviewed and approved by clinical pharmacologists and toxicologists [12]. Any disagreement will prompt a review of the vignette. Final vignette approval will only be possible if a consensus can be reached.
### Dose limit for a single LA (local anesthetic)
- Levobupivacaine: 2 mg/kg (maximum 150 mg/dose)
- Lidocaine: 3 mg/kg (maximum 300 mg/dose)
- Ropivacaine: 3 mg/kg (maximum 225 mg/dose)

### Influence of epinephrine on dose limit
- Levobupivacaine: 3 mg/kg (maximum 150 mg/dose)
- Lidocaine: 7 mg/kg (maximum 400 mg/dose)
- Ropivacaine: 3 mg/kg (maximum 225 mg/dose)

### Determination of calculation weight (CW)
- Calculation of BMI
- Calculation of ideal body weight (IBW; Devine formula)
- Application of the following algorithm to define CW:
  - weight \leq 70 kg and BMI < 30 and IBW > weight → CW = weight
  - weight > 70 kg and BMI < 30 and IBW \leq weight → CW = IBW
  - weight \leq 70 kg and BMI \geq 30 → CW = IBW
  - weight > 70 kg and IBW > 70 → CW = 70
  - weight > 70 kg and IBW \leq 70 → CW = IBW

### Dose adaptation depending on health conditions and drugs
- Conditions
  - Old age (70 years or older)
  - Renal dysfunction (glomerular filtration rate [GFR] < 50 mL/minute)
  - Hepatic insufficiency (prothrombin time < 50%)
  - Heart failure (left ventricular ejection fraction \leq 30%)
  - Pregnancy
  - Drugs decreasing LA metabolism
- List of drugs decreasing LA metabolism
  - Major CYP1A2 inhibitors: ciprofloxacin, norfloxacin, and fluvoxamine
  - Major CYP3A inhibitors: azole antifungals, macrolides, calcium channel blockers, HIV antiretroviral therapy, and tyrosine kinase inhibitors
- If 1 condition is present, the calculator reduces the total maximum dose by 20%
- If 2 or more conditions are present, the calculator reduces the total maximum dose by 30%.

### Calculation rule for LA mixtures
- The app performs the following steps
  - Calculation of maximum safe volume for first LA
  - The user enters which volume of first LA is to be used (0-maximum volume)
  - Calculation of corresponding maximum dose of first LA and determination of percentage of total maximum dose
  - Calculation of maximum dose of second LA based on remaining percentage of total maximum dose
  - Calculation of maximum volume of second LA
Groups and Randomization

There will be 2 study groups: in the control group, participants will be asked to use the method they usually use in their clinical practice to calculate the maximum safe dose of LA; in the LoAD Calc (experiment) group, participants will be required to use the LoAD Calc app, which will be preinstalled on a standard smartphone. The method used by the participants allocated to the control group to calculate the maximum safe dose of LA will be recorded. There will be no teaching or introductory intervention for any of the participants before the study, and the participants allocated to the LoAD Calc group will therefore discover the app while answering the first vignette.

Web-Based Study Platform

A specific web-based platform will be developed using the Joomla! 4.3 content management system (Open Source Matters). It will be hosted on a Swiss server (Kreativmedia GmbH) and secured by the RSFirewall 3 (RSJoomla) and AdminTools 7 (Akeeba Ltd) components. To ensure participant anonymity, unique usernames and passwords will be created using Manytools’ web-based password generator [17]. These credentials will then be imported into Stata and allocated to either study group according to the randomization process described above. Finally, this data will be exported to a CSV file, which will be imported into the web-based study platform using the Import Joomla Users component (version 3.4; Lerus Ltd).

Consents, questionnaires, and vignettes will be managed using Shondalai’s Community Surveys 6 and Community Quiz 6 components (Bulasikku Technologies Pvt Ltd). All data will be stored on an encrypted MySQL-compatible database (MariaDB 10, MariaDB Foundation).

Inclusion and Exclusion Criteria

All resident physicians and registrars working in the HUG anesthesiology department will be eligible for inclusion. The only exclusion criteria will be current or previous use of the LoAD Calc app. This criterion will be assessed by a screening question asked after the completion of all study vignettes.

Recruitment

The project will first be presented to the head of the anesthesiology department and then to all consultants. After obtaining their agreement, investigators will recruit potential participants directly in the operating room. These residents and fellows in anesthesiology will be informed that the study will last at most 1 hour and that an investigator will replace them in the operating room while they participate. They will be told that participation is entirely voluntary, that there will be no consequence if they refuse to participate, and that they will be able to withdraw at any time without explanation. No incentive other than advancing scientific knowledge will be given to promote participation. They will be given a paper sheet summarizing the information regarding the study and data protection (Multimedia Appendix 2). Those who agree will be scheduled for participation on the same day. Together with the anesthesiology consultant overseeing the operating room, an investigator will organize replacements to avoid any disruption in the operating program. There will be only 1 slot, and therefore, only 1 participant per hour.

Consent and Study Sequence

Participants will be asked to set their phones to flight mode. This will enable them to access any note, calculator, or app they use to calculate LA doses while avoiding potentially disruptive interruptions. The study itself will take place in a separate, quiet room. There, an investigator will prompt them to pick up a sealed, opaque envelope containing the credentials necessary to log in.

All participants will be asked to sign an electronic consent form immediately after logging in. Those who agree will proceed to a first questionnaire designed to gather demographical data (Textbox 3) and determine whether these participants are currently using LoAD Calc or if they have used this app before (exclusion criterion). After completing this questionnaire, an introductory screen giving information regarding the vignettes they are about to see and specifying the calculation method they are to use (LoAD Calc for the experimental group vs left at the participant’s will for the control group) will be displayed. At

Textbox 2. Randomization code.

```
set obs #N
egen arm = seq(), to(2)
set seed #S
gen random = uniform()
sort random
```

Wherein “1” will be the control group, and “2” will be the LoAD Calc group.

Since randomization will be stratified according to participant position (either resident or registrar), 2 seeds (#S) will be used (07022023 and 20230207).

A sample size calculation will be used to determine the total number of observations. It will be rounded up to the nearest ten to enhance the study power and take into account attrition and potential exclusions. The stratified number of observations (#N) will be computed according to the proportion of potential participants belonging to both eligible positions (residents vs registrars).

The method used by the participants allocated to the control group to calculate the maximum safe dose of LA will be recorded. There will be no teaching or introductory intervention for any of the participants before the study, and the participants allocated to the LoAD Calc group will therefore discover the app while answering the first vignette.

```
set obs #N
gen random = uniform()
sort random
```

This will enable them to access any note, calculator, or app they use to calculate LA doses while avoiding potentially disruptive interruptions. The study itself will take place in a separate, quiet room. There, an investigator will prompt them to pick up a sealed, opaque envelope containing the credentials necessary to log in.

All participants will be asked to sign an electronic consent form immediately after logging in. Those who agree will proceed to a first questionnaire designed to gather demographical data (Textbox 3) and determine whether these participants are currently using LoAD Calc or if they have used this app before (exclusion criterion). After completing this questionnaire, an introductory screen giving information regarding the vignettes they are about to see and specifying the calculation method they are to use (LoAD Calc for the experimental group vs left at the participant’s will for the control group) will be displayed.

https://www.researchprotocols.org/2024/1/e53679
this stage, those allocated to the LoAD Calc group will be given the smartphone preinstalled with the LoAD Calc app.

Textbox 3. First questionnaire.

<table>
<thead>
<tr>
<th>Page 1: consent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consent to participate and to data reuse (multiple-choice questions with only 1 acceptable answer)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Page 2: exclusion criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has heard of LoAD Calc (multiple-choice questions with only 1 acceptable answer)(^a)</td>
</tr>
<tr>
<td>Has installed LoAD Calc (multiple-choice questions with only 1 acceptable answer; branching logic will be used to avoid displaying irrelevant questions)</td>
</tr>
<tr>
<td>Has used LoAD Calc (multiple-choice questions with only 1 acceptable answer; branching logic will be used to avoid displaying irrelevant questions; answering “yes” to either of those questions will lead to participant exclusion)</td>
</tr>
<tr>
<td>Context of LoAD Calc use (multiple-choice question with only 1 acceptable answer; branching logic will be used to avoid displaying irrelevant questions)</td>
</tr>
<tr>
<td>Has LoAD Calc still installed (multiple-choice question with only 1 acceptable answer; branching logic will be used to avoid displaying irrelevant questions)</td>
</tr>
<tr>
<td>Still uses LoAD Calc (multiple-choice questions with only 1 acceptable answer; branching logic will be used to avoid displaying irrelevant questions; answering “yes” to either of those questions will lead to participant exclusion)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Page 3: demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (multiple-choice questions with only 1 acceptable answer)</td>
</tr>
<tr>
<td>Age (free text with regular expression [regex] validation rule)</td>
</tr>
<tr>
<td>Position (multiple-choice questions with only 1 acceptable answer; custom answer accepted)</td>
</tr>
<tr>
<td>Years since graduation (free text with regular expression [regex] validation rule)</td>
</tr>
<tr>
<td>Years of practice in anesthesiology (free text with regular expression [regex] validation rule)</td>
</tr>
<tr>
<td>Specialist diplomas (custom answer accepted; multiple-answer question with more than 1 answer accepted)</td>
</tr>
</tbody>
</table>

After completing the vignettes, the participants allocated to the LoAD Calc group will be asked to complete the French version of the System Usability Scale [18], the translation of which has been validated [19]. Participants allocated to the control group will then be asked which methods they used to calculate LA doses.

Finally, both groups will have to answer a question based on a 10-point Likert scale to assess their confidence as to the method they used to carry out the maximum safe LA dose calculations, from score 1 (absolutely not confident) to score 10 (perfectly confident).

The whole study sequence is summarized in Figure 1.

Figure 1. Study sequence. LoAD Calc: Local Anesthetics Dose Calculator.
Outcomes
The primary outcome will be the overall overdose rate, according to the method used. To assess this outcome, the maximum acceptable dose in milligrams, or milliliters, must therefore be established. While this might seem straightforward at first sight, it is actually rather complex. Indeed, although maximum safe doses are calculated in milligrams, anesthetists administer a volume of LA (the concentration of which can vary) rather than a quantity of LA. Therefore, even though toxicity is related to the quantity (in milligrams) of LA administered, it is clinically more relevant to determine the maximum volume (in milliliters) of LA that can be used for a particular patient. Consequently, after applying the rules described in Textbox 1, a quantity in milligrams will be obtained. It will then be converted in milliliters according to the concentration of the LA used in the vignette. This volume will then be rounded to the inferior integer. To be less conservative, 1 mm will be added to this calculated volume, and the total will represent the maximum acceptable volume. An overdose will be considered any dose higher than this maximum acceptable volume or than its corresponding LA quantity in milligrams.

The secondary outcomes will be the overall overdose rate, considering the simulated patient’s ideal weight and the simulated patient’s actual weight, the overdose rate according to each LA studied, and the overall underdose rate. An underdose will be defined as the maximum acceptable volume minus 20% (or its corresponding LA quantity in milligrams), rounded to the inferior integer. This is an empirical choice since anesthetic underdose can only be determined clinically [20,21].

Other secondary outcomes will be the time taken to complete these calculations, the app’s usability, and the physicians’ confidence in using the method they were allocated to. The app’s usability will be evaluated using the System Usability Scale [18]. Provided that the statistical assumptions are met, factors associated with a higher probability of overdose or underdose will also be assessed.

Statistical Analysis
The sample size calculation and all other statistical analyses will be carried out using Stata (version 17.0 or above). The complete data set will be exported by the webmaster, who will give the study groups codenames before sending the curated data set for statistical analysis. Descriptive statistics will be used to present demographical data. Normality will be assessed graphically, and the Kolmogorov-Smirnov test will be used in cases of doubt. Accordingly, all outcomes will then be computed using either parametric or nonparametric tests. The data acquisition mechanisms will ensure that all data are recorded after each stage. Thus, there should not be any missing data, and there shall be no need for imputation. When LA mixtures are used, participants will be told that 1 anesthetic has already been injected and the dose used has been clearly reported. Thus, they will be asked to determine the maximum safe dose for the second local anesthetic. Multivariable regression will be used to determine an association between specific clinical parameters and the probability of overdose or underdose, provided that all required assumptions are met and that the risk of overfitting is adequately limited. Double-sided P values (P<.05) will be considered significant.

Results
The 10 vignettes necessary to carry out the study were successfully created, and the maximum safe doses were determined. These vignettes and the doses were checked and approved by all authors. The 10 vignettes, as well as their English translation, were presented to peer reviewers but are not publicly available to avoid any potential bias. They will nevertheless be published along with the results paper.

The sample size calculation was performed using Stata (version 17.0). It showed that 46 participants (23 in each group) would be needed to detect a 10% difference with a power of 90%, taking into account an SD of 10%. In line with the above methods, a total of 50 anesthesiologists should therefore be recruited. Since there are 62 residents and 52 registrars in the HUG Anesthesiology Department, a participation rate of 44% (50/114) will be necessary. This participation rate seems achievable with the aforementioned recruitment procedure. If this rate cannot be achieved, other Swiss University hospitals will be contacted, and similar recruitment procedures will be carried out to obtain the required sample size.

The study platform has been successfully created and tested by all coauthors [22]. The data extraction mechanisms have also been successfully checked.

The recruitment will take place once this study protocol has been reviewed and accepted for publication to allow for any necessary adjustments before study inception. The current version of the protocol is 0.9 (October 10, 2023). The published version will be 1.0.

It should be possible to start recruitment during the winter of 2023. This would allow data analysis to take place in spring 2024, and results should be submitted for publication in an international peer-reviewed journal by the end of the same year.

Discussion
Overview
This study should allow us to determine whether LoAD Calc, an mHealth app designed to calculate maximum safe LA doses, is safer and more effective than current clinical practice. Previous studies have shown that mHealth apps can enhance dose calculation and potentially improve safety [23], decrease time to drug delivery [23], and lessen stress [24]. Assessing this latter parameter would not make much sense given the design of this study but could prove interesting in future high-fidelity or field trials.

Other solutions have already been proposed for the calculation of the maximum safe LA dose but present significant drawbacks. Some of them, such as the nomogram created by Williams and Walker [11], do not depend on technological devices. This nomogram, which represents a rapid and calculation-free way, must, however, always be within reach. In addition, IBW must first be determined, and there is no dose adaptation based on health conditions or drug interactions. Computer-based solutions...
and mobile apps have also been created, such as MDCalc Local Anesthetic Dosing Calculator [25], The Podiatry Institute’s LA Toxic Dose Calculator [26], and SafeLocal by Johns Hopkins Digital [27]. All these solutions lack key elements and do not consider either IBW, comorbidities, or medications. Most allow invalid data to be entered or suggest doses exceeding the maximum safe dose, thereby presenting potential safety issues. No study can be devoid of limitations, and the one planned according to this protocol is no exception. The first foreseeable limitation is that the LoAD Calc app will be compared to many different methods of LA dose calculation, thereby preventing us from directly comparing this app to a specific method. However, there is no gold standard to calculate the maximum safe LA doses, and the design of the proposed study can be considered pragmatic. Another limitation is that the maximum safe doses will be calculated using the same scientifically grounded rules that were used to develop LoAD Calc [12]. However, some of the calculation rules used by the app are not supported by strong scientific evidence, and there is no gold standard for comprehensive, safe calculation of maximum LA doses. Finally, the results obtained through this study will only apply to the single-dose administration of a limited number of LAs or of LA mixtures. This will not affect the validity of the study’s results nor compromise the use of the app since the 3 LAs selected (levobupivacaine, lidocaine, and ropivacaine) are commonly used in clinical practice. Nevertheless, further app developments will be needed to take other LAs and repeated doses into account. Since some LAs, such as lidocaine, are also safe for intravenous use, future versions of the app should enable practitioners to select different injection sites and routes.

Conclusions
Following this protocol should enable us to determine whether LoAD Calc, a mHealth app designed to calculate the maximum safe doses of LA, is both safe and effective. If this hypothesis proves to be true, clinical trials could be considered, and further outcomes, such as the impact of LoAD Calc on cognitive load and physiologic stress, could be considered.

Acknowledgments
The authors of this protocol would like to thank all members of the original LoAD Calc development team for their support throughout its development.

Data Availability
The data sets generated during and/or analyzed during this study will be available in an open access repository.

Conflicts of Interest
None declared.

Multimedia Appendix 1
SPIRIT checklist.
[PDF File (Adobe PDF File), 555 KB - resprot_v13i1e53679_app1.pdf ]

Multimedia Appendix 2
Study information sheet.
[PDF File (Adobe PDF File), 84 KB - resprot_v13i1e53679_app2.pdf ]

References


32. Suppan M. Local Anesthetics Study. URL: https://local.anesth.ch/ [accessed 2023-12-13]


Abbreviations

CCER: Commission Cantonale d’Ethique de la Recherche
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HHEalth Applications and onLine TeleHealth
HUG: Geneva University Hospitals
IBW: ideal body weight
LA: local anesthetic
LAST: local anesthetic systemic toxicity
LoAD Calc: Local Anesthetics Dose Calculator
mHealth: mobile health
SPIRIT: Standard Protocol Items: Recommendations for Intervventional Trials

Edited by G Eysenbach; submitted 15.10.23; peer-reviewed by N Buckley, T Girard; comments to author 02.11.23; revised version received 04.11.23; accepted 21.11.23; published 03.01.24.

Please cite as:
Fubini PE, Savoldelli GL, Beckmann TS, Samer CF, Suppan M
Impact of a Mobile App (LoAD Calc) on the Calculation of Maximum Safe Doses of Local Anesthetics: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2024;13:e53679
URL: https://www.researchprotocols.org/2024/1/e53679
doi:10.2196/53679
PMID:38170571

©Pietro Elias Fubini, Georges Louis Savoldelli, Tal Sara Beckmann, Caroline Flora Samer, Mélanie Suppan. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 03.01.2024. 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.
Professional Development to Improve Responsible Beverage Service Training: Formative Research Results and Protocol for a Randomized Controlled Trial

W Gill Woodall1*, PhD; David Buller1*, PhD; Robert Saltz2*, PhD; Lila Martinez1*, BA

1Klein Buendel, Inc, Golden, CO, United States
2Prevention Research Center, Pacific Institute for Research and Evaluation, Berkeley, CA, United States
*all authors contributed equally

Abstract

Background: Improved interventions are needed to reduce the rate of driving while intoxicated. Responsible beverage service (RBS) training has reduced service to intoxicated patrons in licensed premises in several studies. Its efficacy might be improved by increasing the proper application and continued use of RBS with a professional development program in the 3 to 5 years between the required RBS retraining.

Objective: This study aims to develop and evaluate a professional development component for an RBS training that aims to improve the effectiveness of the web-based training alone.

Methods: In a 2-phase project, we are creating a professional development component for alcohol servers after completing an RBS training. The first phase involved formative research on the feasibility, acceptability, and potential effectiveness of components. Semistructured interviews with owners and managers of licensed establishments and focus groups and a survey with alcohol servers in New Mexico and Washington State examined support for RBS and the need for ongoing professional development to support RBS. A prototype of a professional development component, WayToServe Plus, was produced for delivery in social media posts on advanced RBS skills, support from experienced servers, professionalism, and basic management training. The prototype was evaluated in a usability survey and a field pilot study with alcohol servers in California, New Mexico, and Washington State. The second phase of the project will include full production of the professional development component. It will be delivered in Facebook private groups over 12 months and evaluated with a sample of licensed premises (ie, bars and restaurants) in California, New Mexico, and Washington State (n=180) in a 2-group randomized field trial (WayToServe training only vs WayToServe training and WayToServe Plus). Licensed establishments will be assessed for refusal of sales to apparently intoxicated pseudopatrons at baseline and 12 months after the intervention commences.

Results: Although owners and managers (n=10) and alcohol servers (n=43) were favorable toward RBS, they endorsed the need for ongoing support for RBS for servers and identified topics of interest. A prototype with 50 posts was successfully created. Servers felt that it was highly usable and appropriate for themselves and the premises in the usability survey (n=20) and field pilot test (n=110), with 85% (17/20) and 78% (46/59), respectively, saying they would use it. Servers receiving the professional development component had higher self-efficacy (d=0.30) and response efficacy (d=0.38) for RBS compared with untreated controls.

Conclusions: Owners, managers, and servers believed that an ongoing professional development component on RBS would benefit servers and licensed premises. Servers were interested in using such a program, a large majority engaged with the prototype, and servers receiving it improved on theoretic mediators of RBS. Thus, the professional development component may improve RBS training.

Trial Registration: ClinicalTrials.gov NCT05779774; http://tinyurl.com/4mw6d2vk
alcohol; driving while intoxicated; responsible beverage service; training; prevention; professional development; social media

Introduction

Background

Driving while intoxicated (DWI) is one of the most preventable public health risks in the United States. However, from 2019 to 2020, there was an increase of 14.3% in DWI deaths, after remaining largely stable from 2015 to 2019 [1]. Although new policies and interventions are needed to reduce the consequences of DWI, gains are possible by increasing the efficacy of existing interventions. Responsible beverage service (RBS) training [2-4] has been effective in some cases [5], but methods of boosting its efficacy are needed. The goal of this research is to develop and evaluate a professional development component for a web-based RBS training program that aims to improve the effectiveness of web-based training. The professional development component will provide ongoing information and instruction in advanced RBS techniques and emphasize professionalism in the hospitality industry. It will be continuously available and easily accessible to alcohol servers via social media after completing the web-based training. Marketplace approaches to prevent alcohol service that results in intoxication or restrict access to alcohol by persons already intoxicated are an alternative policy approach to deterrence of driving by drinkers considered impaired to decrease DWI [6]. Most US states have laws prohibiting sales of alcohol to visibly intoxicated customers [7]. A complementary intervention is RBS training, which aims to instruct servers on how to prevent intoxication by teaching drink counting techniques, ways to recognize signs of intoxication, and strategies to refuse alcohol sales. This environmental intervention aims to decrease opportunities for risky behavior [8], consistent with harm reduction in a nurturing environment perspective [9]. It is a targeted restriction on alcohol accessibility at the times and places where risk is greatest that does not depend on decision-making by persons considered alcohol impaired, can be applied to all alcohol sales premises, does not depend on house policies of licensees to refuse sales, and reduces intoxicated customers’ ability to shop around to find premises that will serve.

Research on RBS training presents a mixed picture [5]. Although some studies have failed to show effectiveness [10], reviews in 2000 and 2001 concluded that RBS training can prevent alcohol overservice [11] with strong management support [12]. Recent studies have found RBS training to be associated with increases in refusals of service to apparently intoxicated customers. In addition it was related to decreases in blood alcohol concentration and calls to emergency services [13]. RBS training combined with enforcement reduced alcohol overservice and violent assaults in a trial in Sweden [14-16] but not in Norway [17]. In addition, lower levels of motor vehicle crashes with a high percentage of alcohol involvement were observed in a mandatory RBS training state [18], and another analysis found that states with RBS laws had a reduced number of underage drinking driver fatality crash ratios [19]. Data on self-reported DWI have been mixed, with one study showing no association [20] and another finding a decrease in reported DWI with RBS training [13]. Continued research on RBS training [21] is warranted because (1) positive outcomes have been reported [5,16,22,23]; (2) methodological problems limit existing evidence (eg, lack of randomized trials, clear outcome variables, training fidelity data, and effect size reporting) [5,11,12]; and (3) data are limited on web-based training that can improve training engagement, fidelity, and quality, compared with in-person training [5]. Our team showed that a web-based RBS training program, named WayToServe, was effective in premises serving alcohol for onsite consumption (ie, bars and restaurants) [24].

RBS training laws are highly variable across US states [25]. It is legally mandated in 25 US states and incentivized in some fashion in a number of other states [26]. Most states that require training have long periods of 3 to 5 years between required retraining. Consequently, the proper application, monitoring, and continued use of the RBS techniques falls on the shoulders of premises management, so it is not surprising that management commitment to RBS can affect servers’ adherence to RBS methods [5,12]. Developing ways to support RBS techniques after training may counter the management’s ineffective or limited support for, disinterest in, or outright resistance to RBS.

Objectives

The goal of this study is to develop and evaluate a follow-on professional development component to increase the efficacy of our web-based RBS training, WayToServe. Continuing professional development is a widespread practice across a variety of fields including accounting, social work, and medicine [27]. Typically, it focuses on improving knowledge, skills, and performance to help employees stay up-to-date on industry developments, develop and maintain job capabilities, convey professional values and norms, and create communities of practice [27-32]. Although often focused on high-skill professional workers (eg, nurses, physicians, lawyers, and architects) [27], the training and certification of community members has improved professionalism [33] and, along with in-service contact, has boosted the success of community prevention programs [34,35]. Vocational education and lifelong learning play essential roles in the hospitality industry. They offer both general knowledge and skills such as communication and customer service as well as job-related knowledge such as understanding laws related to serving alcohol and the ability to recognize signs of intoxication [32,36,37]. Alcohol servers trained in RBS practices should benefit from ongoing professional development by (1) motivating them to implement RBS skills in the face of common barriers, such as pressure to...
sell, low management support, and customers’ attempts to continue being served; (2) receiving support for RBS from a community of alcohol servers, especially for servers who work in small or unsupportive premises; and (3) preventing the degradation of RBS skills over time.

**Methods**

**Overview**

This study is being conducted in 2 phases to create an effective professional development component for alcohol servers who completed state-approved RBS training. The first phase aimed to determine whether a professional development component delivered over social media was feasible, acceptable, and potentially effective for alcohol premises management and servers. The second phase will involve production of the professional development component and testing its efficacy in a randomized controlled trial (RCT).

**Phase 1: Formative Research**

**Overview**

A formative research phase in 2022 used both qualitative and quantitative techniques to provide a detailed picture of RBS training and its gaps as well as the feasibility and potential effectiveness of adding a professional development component. It included semistructured interviews with premises managers, focus groups and a survey with alcohol servers, development of a prototype of the professional development component, and usability testing and a field pilot study evaluating the prototype.

**Semistructured Interviews With Alcohol Premises Owners and Managers**

Owners and managers (n=10) of onsite alcohol premises in New Mexico and Washington State participated in semistructured interviews. Premises had to hold an active state alcohol license and be a bar or restaurant that served alcohol. Managers discussed RBS policies at their premises, support for RBS methods, perceived importance of RBS, quality of RBS methods implemented by their servers, and need for ongoing training and support for RBS among their servers. They also suggested content for a professional development component. The transcripts were coded to identify themes.

**Focus Groups and Survey With Alcohol Servers**

Alcohol servers working in onsite alcohol sales premises in New Mexico and Washington State were recruited from the roster of WayToServe trainees to provide input on experience with RBS methods and interest in ongoing professional development related to RBS methods. To be included, servers had to be aged ≥19 years (by state regulation), serve alcoholic beverages at a licensed premises, have completed the WayToServe RBS training, and be proficient in English. Initially, alcohol servers (n=19) were recruited to web-based focus groups; however, when participation lagged, servers (n=24) were recruited instead to complete a web survey. In both the focus groups and survey, servers were asked about their experience with, confidence in, and barriers to RBS methods; support from management and other alcohol servers for RBS; and experience refusing service to customers. They also indicated their interest in and potential utility of receiving ongoing information and activities from WayToServe to keep up-to-date and be capable of using RBS methods via a Facebook group. Servers reported whether they were willing to share their RBS experiences or provide feedback on RBS actions with other servers. Focus group transcripts were coded for themes for each question. The survey responses were summarized using descriptive statistics.

**Production of a Prototype Professional Development Component**

A prototype of the professional development component was produced by the project staff and media developers. Named WayToServe Plus, it comprised a series of 50 social media posts. The goal of the messages was to improve servers’ professionalism by (1) increasing the confidence and motivation of RBS-trained servers to implement RBS methods, with attention to ways of overcoming common barriers; (2) creating a professional community of servers that supports one another in implementing RBS actions and serves as a resource for advice and strategies to implement RBS (eg, tips and tricks) by encouraging servers to share their personal experiences through comments and posts; and (3) preventing the deterioration of RBS skills and motivation over time by providing refresher instruction. Posts contained text, graphics, web-based learning activities, and videos demonstrating RBS techniques in 4 topic areas derived from the results of the manager interviews and alcohol server focus groups and surveys: advanced RBS skills training (ID checking, cannabis and alcohol, drink counting, and home delivery), experienced servers supporting new servers (eg, tips and tricks to apply RBS and sharing stories on RBS experiences), professionalism (safety and security, security personnel, and handling disruptive customers), and basic management training (content and development of house RBS policies and best practices for RBS). Instructional goals included improving the application of RBS information and skills in realistic settings and circumstances that servers have encountered in their jobs. WayToServe Plus was consistent with the transformative approach to continuing professional development by Kennedy [38], combining the transmission of information, skills, and norms and providing coaching or mentoring by striving to create a community of practice among servers with varying levels of experience. Messages in the posts were guided by principles of diffusion of innovation theory (eg, compatibility, simplicity, trialability, and observability) [39] and social cognitive theory (ie, modeling) [40] and written to be relatable, positive, and entertaining. A total of 14 short videos were produced using the TikTok video authoring platform. An interactive quiz activity was taken from the WayToServe training and linked to a social media post. WayToServe Plus was authored in English because most servers had elected to complete WayToServe training in English.

**Evaluation of Prototype Professional Development Component**

The prototype WayToServe Plus component was evaluated for usability, feasibility, acceptability, and engagement through a survey and a field pilot test with alcohol servers recruited from the roster of WayToServe trainees.
Usability Testing
Alcohol servers (n=20) who met the aforementioned inclusion criteria (refer to the Focus Groups and Survey With Alcohol Servers section) completed a web survey on prototype acceptability, feasibility, and utility (10 in New Mexico and 10 in Washington State). Ten usability testers can identify 95% of problems [41-43]. In the survey, servers were provided with a description of the WayToServe Plus component. Each server was presented with 3 posts and 1 video randomly selected from posts in the prototype. They were then asked to evaluate these items based on their appropriateness for themselves and licensed establishments, their acceptability, and usefulness, using 5-point Likert scales. Servers indicated if they would read or view the post or video, react to it (eg, like, sad, and angry), comment on it, and share it. A description of the interactive activities was provided and rated based on these measures. Next, servers evaluated the WayToServe Plus concept on the validated System Usability Scale (SUS) [44-46]. The 10 items were combined using standard techniques, with a score of ≥68 indicating adequate usability. In addition, a single item assessed user-friendliness (1=worst imaginable and 7=best imaginable). Finally, servers indicated whether they would be interested in getting the ongoing information and activities from WayToServe Plus, topics that would be of interest to them, and potential reasons for not using it.

Field Pilot Test
A sample of 59 alcohol servers (5 in California, 21 in New Mexico, and 33 in Washington State) participated in a 1-month field pilot test of the prototype, meeting the same inclusion criteria as the focus group and usability testing participants. The study involved a nonrandomized posttest-only 2-group design, in which the treatment group had 2 levels: prototype WayToServe Plus program versus no treatment. In the intervention group, 59 servers were recruited and joined a Facebook private group on a rolling basis over 8 weeks. Staff posted prototype WayToServe Plus posts (1 per day, Monday to Friday) for the 8-week period. Approximately 24 posts were posted to the private group during any 4-week period in the intervention period. The posts were only viewable to members of the private group and could not be shared outside the private group. Outcomes were assessed in 2 ways at 4 weeks after enrollment. First, servers’ engagement with WayToServe Plus was measured by recording the number of times posts were viewed, reacted to (eg, liked), and commented on by servers. Second, servers completed a web-based posttest, assessing the prototype on appropriateness, acceptability, and utility for servers and premises and usability on the SUS [44-46] and whether the tone of the prototype aligned with their licensed establishment’s atmosphere, using scales similar to those used in the usability survey. Perceived self-efficacy and response efficacy for maintaining community safety by using RBS methods (5-point Likert scales) were measured as proxy outcomes of the effectiveness of the prototype program (ie, dependent variables). Willingness to use the WayToServe Plus program in the future and job and demographic characteristics were also measured. The respondents suggested ways to improve the prototype and make it more engaging. A second group of 51 servers was recruited to serve as an untreated control group and completed only a posttest web survey, with the primary purpose being to assess their perceived self-efficacy and response efficacy of RBS methods and compare them with ratings provided by servers who received the prototype.

Phase 2: RCT Protocol for Evaluating the Professional Development Component
The WayToServe Plus professional development component will be fully produced and evaluated for effectiveness in an RCT.

Production and Implementation of WayToServe Plus Component
A 12-month version of the WayToServe Plus professional development component will be created for evaluation in the trial. Content and format will be developed according to instructional goals, principles from diffusion of innovation theory and social cognitive theory, and insights derived from the formative research findings in phase 1. Posts will contain text, infographics, short videos, and interactive activities based on the WayToServe RBS training. Features to elicit user-generated content will be included in posts, such as posing a common situation and asking, for example, RBS strategies; providing polls about RBS methods; and soliciting stories, tips, and tricks from experienced servers for applying RBS. These posts are intended to create sense of community among alcohol servers. An agile iterative production process will be used to author the posts [47]. Approximately 2 months of posts will be prepared before launching the intervention; additional posts will be developed during the intervention, adjusting them for season, current events, and reactions and comments from servers to prior posts.

WayToServe Plus will be administered by a staff member who serves as a community manager. The manager will post 4 posts per week, Monday through Friday, for 12 months (approximately 208 posts in total). In addition, posts selected from the usual-and-customary WayToServe Facebook page will be posted once per week. Alcohol servers can comment on and react to posts but cannot share them on their own feed. Orientation to private groups will be self-explanatory. The community manager will stress respect for others; monitor comments; and correct inappropriate, unfavorable, or bullying comments or misinformation [48]. Servers will be compensated US $50 for joining the WayToServe Plus Facebook private group.

Randomized Trial Design
The WayToServe Plus professional development component will be evaluated with a sample of 180 establishments licensed for sale of alcohol for onsite consumption (ie, liquor by the drink) and their alcohol servers. Premises will be enrolled in a 2-group randomized field trial (WayToServe training only [comparison control] vs WayToServe training plus WayToServe Plus [intervention]) with 2 assessment rounds (baseline and posttest [12 months after intervention commences]), yielding a 2 (treatment)×2 (assessment time) factorial design. Using a custom-written program, the project biostatistician will randomly assign half of the premises (90/180, 50%) to WayToServe training plus WayToServe Plus, stratified by state (ie,
Selection and Recruitment of Licensed Alcohol Premises

State-licensed onsite alcohol establishments (n=180) in California (n=60), New Mexico (n=60), and Washington State (n=60) were randomly selected from publicly available lists from state alcohol regulation agencies, stratified by location (metro areas [Albuquerque, San Francisco [including Oakland and San Jose], and Seattle; n=148 premises] vs suburban towns [n=7 towns and 32 premises]). As in the formative research, they had to hold an active state license to sell alcohol and be a bar or restaurant that sold alcoholic beverages. To control travel costs in the large San Francisco and Seattle metropolitan areas, clusters of establishments were constructed by randomly selecting seed premises. Next, 14 additional establishments were randomly selected from within the same zip codes of the seed premises. Within each seed area, 5 to 7 establishments were selected at random for PiP visits, with the remaining premises serving as replacements for any deemed ineligible (eg, do not sell alcohol for onsite consumption) or that were closed (either permanently or during evening hours) when visited by PiP teams. In New Mexico, the Albuquerque metropolitan area was much smaller geographically, so we selected premises at random from the state lists.

After the baseline PiP assessment, project staff will contact premises management, describe participation, record the number of alcohol servers, and obtain agreement to participate in the study. Premises will be given a voucher to provide to their servers to access the WayToServe training and complete it within 4 weeks from registration. Servers will complete a consent form. For completing the training, servers will receive US $35 and a new server training certificate for their state. WayToServe will remain available to the participating premises throughout the trial, and managers will be asked to have newly hired alcohol servers complete it.

PiP Assessment Protocol

The primary outcome (ie, dependent variable) will be refusal of sale of alcoholic beverages to visibly intoxicated patrons assessed using a PiP protocol. Ethnically diverse male and female legal-age individuals (aged ≥ 21 years) will be hired as confederates, chosen for prior acting experience, and trained to feign intoxication when acting as buyers [50-52]. Signs of intoxication (ie, fumbling with keys or cash, swaying, slurred speech, and stumbling) indicate a high level of alcohol intoxication [53], provide a clear unambiguous choice whether to serve, and are signs that alcohol servers are trained to recognize in the WayToServe training and WayToServe Plus component. In each round (baseline and posttest), assessment will involve 2 PiP buyer visits per premises by the PiP team comprising a buyer and an observer, separated by at least 6 weeks. At each visit, observers will enter the premises before the buyer and position themselves to be able to see the buyer-server interactions. Buyers will enter the premises displaying intoxication signs and order an inexpensive beer. Both buyers and observers will record if alcohol servers agree to serve the buyers the requested alcoholic beverage. Buyers will also record if the drink was served either as requested, with reluctance, with a joke or similar remark, or with a warning that no future drinks will be served.

In addition, buyers will note the type of beverages requested, if their ID was requested, and other responses by the alcohol servers (made statements of risk, enlisted other patrons to support nonsale, offered a nonalcoholic beverage instead, offered food, provided other information [offer of taxi or safe ride, drinking facts, etc], or delayed or ignored service). Observers will record the characteristics of the establishments (state, type, number of staff and patrons, warning signs posted, and cleanliness), rate how busy the establishment is and speed of service, note if staff appear overly familiar with customers, and record the behavior of buyers (type of drink ordered, signs of intoxication displayed, and rating of obviousness of signs of intoxication). Both buyers and observers will record the servers’ job at the establishment (bartender, server, manager, bouncer, or other) and apparent sex (male, female, or do not know), Hispanic ethnicity, and race.

Outcome Analysis

The analysis of study outcomes will test the following hypothesis that compared with premises in WayToServe RBS training only group, premises assigned to receive WayToServe RBS training and WayToServe Plus will have higher rates of refusing PiP at posttest.

In our prior research, the uptake of training in alcohol establishments affected refusal rates [54], so we will test whether improvements in refusal rates are associated with uptake of the WayToServe training and engagement with the WayToServe Plus component. Training uptake will be obtained from the WayToServe web-based program database (ie, the number of servers registered, training modules completed, and completion of the training). Engagement with WayToServe Plus will be assessed by counting the number of reactions and comments on posts by servers within each premise [55]. We will not be able to count the views of posts because our sample size exceeds...
250 participants; Facebook does not report views of posts in private groups with >250 users. Characteristics of alcohol establishments (type of license, type of business [bar or restaurant], how busy the premises was, and number of staff present during visit), alcohol servers interacting with PiP buyers (sex and ethnicity observed by PiP observers), and PiP buyers (sex and ethnicity) will be analyzed initially as control variables and then in subsequent models as effect modifiers of WayToServe Plus.

**Interviews of Owners and Managers on WayToServe Plus Feasibility**

After posttesting of establishments is completed, 18 owners and managers (6 per state) from premises in the WayToServe training and WayToServe Plus groups will be selected at random for interviews about WayToServe Plus, its compatibility with premises’ RBS policies and practices, helpful features, server engagement, suggested improvements, and problems or barriers (compensation=US $75). In addition, alcohol servers in these premises will be surveyed about the same issues and report their engagement with WayToServe Plus, whereas servers in the control premises will be surveyed about the WayToServe web-based training.

**Ethical Considerations**

The protocols used in the formative research and randomized trial were reviewed and approved by the WCG institutional review board (#20211770). Participants read and signed an informed consent form (interviews, focus groups, and pilot field trial) or read and acknowledged a consent statement (surveys) approved by the institutional review board that described the purpose of the research, the research procedures, known risks and benefits, and the use and security of the data. Participants were informed that their participation was voluntary and that their identity would not be disclosed in any public presentation. The participants were compensated as follows: interviews (US $75), focus group discussion (US $50), survey (US $25), usability test (US $50), and field pilot study (US $100).

**Results**

**Phase 1**

**Owner and Manager Interviews About Professional Development**

All owners and managers (n=10; 4 female individuals; 1 African American and 5 Hispanic White individuals) indicated that they supported RBS methods to maintain their establishment’s reputation in the community; keep the community safe; and avoid fines, disruptions, and other problems. They supported alcohol servers by addressing RBS methods in mandatory staff meetings and trainings and manager logs. They said that support for RBS was provided by experienced staff. Although most felt the RBS methods were effective at their premises, they did indicate that bartenders have many job tasks and need more help, and some establishments had more difficulty with RBS methods during summers when patrons drank longer and larger quantities of alcohol.

All owners and managers endorsed the need for ongoing training and support for RBS methods and felt that a program could help them support RBS practices. They desired topics such as checking IDs, new recreational marijuana laws, special venues (eg, music venues, wineries, and events), and communication and conflict resolution. They preferred formats such as educational memes, videos, shared experiences and tips and tricks from experienced servers, resource pages, reminders, and work group chats. Owners and managers felt that a variety of staff would benefit from ongoing training. Some of them did not feel confident addressing topics such as marijuana laws, how to handle patrons with children, and how to manage servers’ desires to sell alcoholic beverages and make money. All owners and managers would be interested in a program that provided ongoing training and support for RBS if it provided new, relevant content in engaging, easily digestible formats without a large time commitment. They were mostly or very likely to use such a program with their alcohol service staff.

**Focus Groups and Survey of Alcohol Servers About Professional Development**

**Focus Groups**

Alcohol servers participating in the focus groups (n=19) were employed in bars, restaurants, and other premises (eg, ski resort, theatre, and market) as bartenders, servers, and other staff. They had worked as servers for 2 months to 6 years.

Most servers had positive experiences applying RBS methods. They reported that owners and managers at their establishment considered RBS methods to be positive, took them seriously, and supported using them. A few servers said that they received very little support, support only from direct supervisors, and support only when they did something incorrectly. They cited management turnover and very large venues as situations that reduced support for RBS. Obstacles to RBS included customers drinking before arriving; pressure to not check IDs or provide heavier pours to regular customers or members of clubs; customers not wanting to hand over ID during COVID-19 social distancing rules; ability to check IDs from different states or military IDs; fake IDs; pressure to sell and fear of losing tips; large, busy events; and potential for negative reactions when refusing sales to intoxicated patrons. Methods to overcome obstacles included observing customers when they arrive; relying on managers, bartenders, and other servers for help; slowing down activities during busy times or slowing service to intoxicated customers; eliminating tips; having support from management; dividing RBS tasks among different staff; serving water to intoxicated customers; and setting limits on the number of drinks served.

All servers saw benefits of continued information, training, and support for RBS methods. The servers also described topics that would be helpful. The perceived benefits of a continuing professional development program included peer support, sense of community, networking among servers, keeping updated on new information, and helping new servers. These included interactive learning activities, prompts, refreshers on laws, IDs,
recognizing intoxication, drink counting, and refusing service, instruction on how to deal with minors, forums for sharing stories and tips with other servers, help for newer servers, ways of managing difficult customers and de-escalating conflict, polls, infographics, and reminders. Almost all servers indicated that they would be willing to share their experiences and provide feedback on RBS methods in the WayToServe Plus Facebook group.

**Survey**

Table 1 presents the profile of the sample of alcohol servers (n=24) who completed the web survey. Alcohol servers ranged in age from 20 to 40 (mean 28.8, SD 5.4) years; a majority were non-Hispanic White (7/24, 29% Hispanic), and there were slightly more female individuals than male individuals. Most servers worked in restaurants and bars as alcohol servers and bartenders and were a mix of new (7/24, 29% had worked less than 1 year) and experienced (15/24, 63% had worked 3 years or more) servers.

Although most servers were very sure of their ability to apply RBS methods, some encountered problems. A sizable minority were only somewhat sure or unsure that they could verify the validity of IDs (3/24, 12%; mean 4.79 out of 5, SD 0.66), check IDs for age of patron (2/24, 8%; mean 4.83, SD 0.64), count number of drinks to prevent intoxication (6/24, 25%; mean 4.63, SD 0.77), recognize if a patron is intoxicated (6/24, 25%; mean 4.63, SD 0.88), or refuse alcohol service to an intoxicated patron (5/24, 21%; mean 4.75, SD 0.53). The obstacles to RBS methods cited included busy serving environments (10/24, 42%), customer intoxicated before arriving (10/24, 42%), regular customers expecting heavier pours (6/24, 25%), coworker or management pressure to not follow RBS regulations (5/24, 21%), fear of sacrificing tips (3/24, 12%), and checking IDs in a group of customers (2/24, 8%). Servers suggested several ways to overcome these obstacles, such as checking everyone’s ID (11/24, 46%), getting support from management to follow rules (5/24, 21%), taking a moment to breathe in busy environments (9/24, 38%), serving water (6/24, 25%) or food (5/24, 21%) to patrons that need to sober up, monitoring patrons for signs of intoxication (13/24, 54%), and involving a manager in handling difficult customers (13/24, 54%).

Almost all servers felt that the management of their establishment was supportive of RBS methods, agreeing that management believes RBS methods are beneficial (22/24, 92%; mean 4.54 out of 5, SD 0.66) and management takes RBS methods seriously (22/24, 92%; mean 4.67, SD 0.64). However, 21% (5/24) reported that management at their establishment provided support for RBS methods only sometimes, rarely, or never. The most common support was help serving when the establishment gets busy (16/20, 80%), answering questions about RBS methods (17/21, 81%), helping servers refuse service to a customer (14/20, 70%), helping check IDs during busy periods (9/20, 45%), and highlighting things to be on alert for before a shift (9/19, 47%).

Servers favorably evaluated the idea of professional development. Overall, 71% (17/24) of servers expressed interest in receiving ongoing information and activities from WayToServe to help keep them up-to-date and be able to use the RBS methods. The benefits servers saw from this professional development for themselves would be receiving tips and tricks from other servers, getting refreshers on everyday work practices, helping other servers who need it, and providing a place to vent about poor experiences while serving alcohol (Table 2). Benefits for the establishment included having servers be on the same page when it comes to serving alcohol and remaining in good standing with the state’s alcohol licensing agency. Finally, 46% (11/24) of the servers said they were somewhat or very likely to join a Facebook group with the professional development content, and 21% (5/24) might join it. The topics of most interest to servers included refreshers on signs of intoxication, unusual or humorous experiences by another server, quizzes that test knowledge of alcohol serving laws, servers sharing positive or negative on-the-job experiences, stories from other servers about how they used an RBS method, and polls on what the servers believe the community thinks about alcohol serving topics (Table 2). Less popular topics were servers sharing experiences via Facebook Live, refreshers on laws and penalties, information on new state laws, refreshers on ID checking, interactive learning activities, and instruction on using RBS methods.
Table 1. Profile of alcohol server samples in formative research.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Server survey (n=24)</th>
<th>Usability testing (n=20)</th>
<th>Field pilot study</th>
<th>Prototype group (n=59)</th>
<th>Control group (n=51)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Type of licensed sales, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On site (by the drink)</td>
<td>__a</td>
<td>13/20 (65)</td>
<td>38/59 (64)</td>
<td>39/51 (77)</td>
<td></td>
</tr>
<tr>
<td>Off site (package)</td>
<td>1/20 (5)</td>
<td>7/59 (12)</td>
<td>8/51 (16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>6/20 (30)</td>
<td>13/59 (22)</td>
<td>4/51 (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Type of establishment, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bar</td>
<td>4/24 (17)</td>
<td>1/20 (5)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Restaurant</td>
<td>12/24 (50)</td>
<td>14/20 (70)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Nightclub</td>
<td>0/24 (0)</td>
<td>2/20 (10)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Brewery</td>
<td>3/24 (12)</td>
<td>2/20 (10)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Distillery or winery tasting room</td>
<td>0/24 (0)</td>
<td>1/20 (5)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5/24 (21)</td>
<td>1/20 (5)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td><strong>Job type, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartender</td>
<td>8/24 (33)</td>
<td>7/20 (35)</td>
<td>15/59 (25)</td>
<td>13/51 (26)</td>
<td></td>
</tr>
<tr>
<td>Server</td>
<td>12/24 (50)</td>
<td>9/20 (45)</td>
<td>27/59 (46)</td>
<td>24/51 (47)</td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>2/24 (8)</td>
<td>1/20 (5)</td>
<td>7/59 (12)</td>
<td>7/51 (14)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2/24 (8)</td>
<td>3/20 (15)</td>
<td>9/59 (15)</td>
<td>6/51 (12)</td>
<td></td>
</tr>
<tr>
<td><strong>Years of experience, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>7/24 (29)</td>
<td>1/20 (5)</td>
<td>14/59 (24)</td>
<td>6/51 (12)</td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>1/24 (4)</td>
<td>1/20 (5)</td>
<td>14/59 (24)</td>
<td>9/51 (18)</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>8/24 (33)</td>
<td>5/20 (25)</td>
<td>10/59 (17)</td>
<td>14/51 (27)</td>
<td></td>
</tr>
<tr>
<td>&gt;5</td>
<td>8/24 (33)</td>
<td>13/20 (65)</td>
<td>20/59 (34)</td>
<td>22/51 (43)</td>
<td></td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>28.8 (5.3)</td>
<td>32.2 (4.7)</td>
<td>33.1 (11.5)</td>
<td>34.5 (11.7)</td>
<td></td>
</tr>
<tr>
<td><strong>Race or ethnicity, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0/24 (0)</td>
<td>0/20 (0)</td>
<td>1/51 (2)</td>
<td>0/51 (0)</td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1/24 (4)</td>
<td>2/20 (10)</td>
<td>2/51 (4)</td>
<td>3/51 (6)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0/24 (0)</td>
<td>0/20 (0)</td>
<td>1/51 (2)</td>
<td>4/51 (8)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>7/24 (29)</td>
<td>8/20 (40)</td>
<td>13/59 (22)</td>
<td>17/51 (33)</td>
<td></td>
</tr>
<tr>
<td>Native Hawaiian or other Pacific Islander</td>
<td>2/24 (8)</td>
<td>0/20 (0)</td>
<td>4/51 (8)</td>
<td>0/51 (0)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>11/18 (61)</td>
<td>9/20 (45)</td>
<td>27/48 (56)</td>
<td>30/51 (59)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex, n/N (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>13/24 (54)</td>
<td>12/19 (63)</td>
<td>41/59 (70)</td>
<td>18/48 (38)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>10/24 (42)</td>
<td>7/19 (37)</td>
<td>17/59 (30)</td>
<td>29/48 (60)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1/24 (4)</td>
<td>0/19 (0)</td>
<td>0/59 (0)</td>
<td>1/48 (2)</td>
<td></td>
</tr>
</tbody>
</table>

*aNot available.*
Table 2. Topics of interest in continuing professional development identified by alcohol servers (n=23).

<table>
<thead>
<tr>
<th>Topic</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benefits to servers</strong></td>
<td></td>
</tr>
<tr>
<td>Tips and tricks from other servers</td>
<td>18 (78)</td>
</tr>
<tr>
<td>Refreshers on everyday work practices</td>
<td>13 (56)</td>
</tr>
<tr>
<td>Providing help to other servers who need it</td>
<td>12 (52)</td>
</tr>
<tr>
<td>Having a place to vent about poor experiences while serving alcohol</td>
<td>11 (48)</td>
</tr>
<tr>
<td><strong>Benefits to establishments</strong></td>
<td></td>
</tr>
<tr>
<td>Having servers on the same page when it comes to serving alcohol</td>
<td>21 (91)</td>
</tr>
<tr>
<td>Remaining in good standing with the state’s licensing body</td>
<td>13 (56)</td>
</tr>
<tr>
<td><strong>Professional development topics of interest</strong></td>
<td></td>
</tr>
<tr>
<td>Stories from other servers about how they used an RBS(^a) method</td>
<td>8 (35)</td>
</tr>
<tr>
<td>Unusual or humorous experience by another server</td>
<td>11 (48)</td>
</tr>
<tr>
<td>Refreshers on signs of intoxication</td>
<td>12 (52)</td>
</tr>
<tr>
<td>Quizzes that test knowledge of alcohol serving laws with prizes</td>
<td>9 (39)</td>
</tr>
<tr>
<td>Servers sharing positive or negative on-the-job experiences</td>
<td>9 (39)</td>
</tr>
<tr>
<td>Polls on what the server community thinks about alcohol serving topics</td>
<td>8 (35)</td>
</tr>
<tr>
<td>Interactive activities that help maintain a skill</td>
<td>7 (30)</td>
</tr>
<tr>
<td>Information on new state laws</td>
<td>7 (30)</td>
</tr>
<tr>
<td>Question and answer posts</td>
<td>9 (39)</td>
</tr>
<tr>
<td>Interactive learning activity for applying an RBS method with feedback</td>
<td>6 (26)</td>
</tr>
<tr>
<td>How to use an RBS method</td>
<td>5 (22)</td>
</tr>
<tr>
<td>Refreshers on ID checking</td>
<td>5 (22)</td>
</tr>
<tr>
<td>Refreshers on laws and penalties pertinent to servers</td>
<td>5 (22)</td>
</tr>
<tr>
<td>Servers sharing their experiences via Facebook Live segments</td>
<td>4 (17)</td>
</tr>
</tbody>
</table>

\(^a\)RBS: responsible beverage service.

**Survey on Acceptability, Feasibility, and Usability of WayToServe Plus Prototype**

Table 1 presents the profile of the sample of alcohol servers in the usability test of the WayToServe Plus prototype (n=20). They were aged 25 to 42 (mean 32.2, SD 4.7) years, and the majority were non-Hispanic White (8/20, 40% were Hispanic) and predominately female individuals. Most worked in on-premises alcohol sales establishments, especially restaurants; however, several worked in nightclubs, breweries, and distillery or winery tasting rooms. The main job types were bartender and alcohol server. Most were experienced alcohol servers, with 90% (18/20) working for >2 years as a server.

Alcohol servers rated the posts in WayToServe Plus prototype as highly appropriate for themselves and their establishment, very acceptable, and useful (with average ratings of all social media posts and the video being above the scale midpoint; Table 3). They evaluated posts on management and house policy most favorably (means ranged from 3.80 to 4.30), compared with posts on additional training (means ranged from 3.30 to 3.85) and disruptive customers (means ranged from 3.58 to 4.05). The prototype videos were very favorably evaluated in terms of appropriateness, acceptability, and usefulness (means 3.80-4.25). Of the 20 servers, 8 (40%) rated the prototype as usable on the SUS, and 90% (18/20) evaluated it as user-friendly (good, excellent, or best imaginable).

Most servers indicated that they would use WayToServe Plus if it was available. Specifically, 60% (12/20) felt that they would like to use it in the future, and 85% (17/20) were interested in getting ongoing information and activities from WayToServe Plus to help keep up-to-date and be able to use RBS methods. When considering specific posts, most servers said they would engage with the posts (view, react to, comment, and share), with the number who would read and react to them being especially high. Slightly fewer servers said they would comment on or share their own posts, but >45% (9/20) said they would do so. Videos were the most engaging, with >70% (14/20) saying they would read, react, comment on, and share them (Table 3). In addition, 85% (17/20) of the servers would use an interactive learning activity if posted in the WayToServe Plus component.
Field Pilot Test of WayToServe Plus

The profile of the alcohol servers participating in the intervention group (n=59) and control group (n=51) in the pilot test is presented in Table 1. Intervention group participants ranged in age from 18 to 65 (mean 33.1, SD 11.5) years, were mostly non-Hispanic White (13/59, 22% Hispanic), and were predominately female individuals. By job type, most worked in on-premises sales establishments as bartenders or alcohol servers. They were a mix of new and experienced servers (14/59, 24% had worked less than 1 year and 20/59, 34% had worked 5 years or more).

Alcohol servers had high engagement with the WayToServe Plus professional development component. Overall, 83% (50/60) viewed at least 1 post, and they viewed an average of 14.85 (SD 12.41) posts over 4 weeks (approximately 24 posts were displayed in any 4-week period). Just under half of the servers (28/60, 47%) reacted (eg, liked) or commented on a post in the WayToServe Plus group. Servers on average reacted to 4.17 posts (SD 7.08).

Alcohol servers evaluated the WayToServe Plus component as highly usable and its content as appropriate. The mean rating on the SUS scale was 81.10 out of 100, with 88% (52/59) giving it a score of ≥68, a common threshold for usability on this scale. They also gave it high marks on user-friendliness (mean 5.81 out of 7). Many felt that the component (52/59, 88% agreed or strongly agreed; mean 4.42 out of 5, SD 1.19) and its content (50/59, 84%; mean 4.31, SD 1.25) were appropriate for them as alcohol servers and aligned with their establishment’s atmosphere (48/59, 81%; mean 4.12, SD 0.79). Most found the posts (49/59, 83%; mean 4.08, SD 0.75) and other servers’ comments on the posts (47/59, 80%; mean 4.05, SD 0.78) to be useful. A large majority of the alcohol servers said that they were likely to use WayToServe Plus in the future (46/59, 78% somewhat likely or very likely; mean 3.95 out of 5, SD 0.99).

Servers in the control group were similar in characteristics to those in the intervention group (Table 1), although control servers had more years of experience on average. Alcohol servers who received the WayToServe Plus prototype were compared with those in the control group in their reported self-efficacy and response efficacy for implementing RBS methods as an indicator of the potential impact of the WayToServe Plus program. Given the small sample size, we planned a priori to calculate the effect size estimate, $d$, rather than perform a standard statistical significance test. Ratings on self-efficacy were higher among servers in the prototype group (mean 4.53, SD 0.57) than servers in the control group (mean 4.33, SD 0.77; $d=0.30$). Likewise, ratings on response efficacy were greater in the prototype group (mean 4.68, SD 0.68) than in the control group (mean 4.39; SD 0.83; $d=0.38$).

### Phase 2

Phase 2 was funded in September 2022. Baseline assessment of licensed alcohol premises (n=179) in California (n=59), New Mexico (n=60), and Washington (n=60) was conducted in 2022-2023 using the pseudopatron protocol, and results are available elsewhere [56]. The recruitment of premises to have servers trained and join the Facebook private group containing the professional development posts is ongoing. Posttest assessment is planned for summer and fall, 2024 with results expected to be published in 2025.

### Discussion

#### Principal Findings

The development of a professional development extension of our RBS training course aims to improve the efficacy of RBS training in the field. The formative research confirmed that owners, managers, and alcohol servers considered a professional development component for RBS to be beneficial, and a large majority would be interested in using such a program. Many owners and managers have already taken steps to help servers implement and maintain their RBS skills, and several of them felt that WayToServe Plus would complement and aid in these efforts. A previous study found that managers trained in RBS also trained their staff in cutting off intoxicated patrons and handling fake IDs [57]. Alcohol servers considered the WayToServe Plus prototype to be highly appropriate, acceptable, usable, and useful. Many servers followed (ie, viewed a post) and engaged (ie, reacted to or commented on a post) with the

<table>
<thead>
<tr>
<th>Table 3. Acceptability and potential engagement with messages in the WayToServe Plus prototype usability survey (N=20).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social media posts by topic</td>
</tr>
<tr>
<td>Acceptability of posts, mean (SD)</td>
</tr>
<tr>
<td>Appropriate for me</td>
</tr>
<tr>
<td>Appropriate for establishment</td>
</tr>
<tr>
<td>Acceptable to me</td>
</tr>
<tr>
<td>Useful to me</td>
</tr>
<tr>
<td>Potential engagement with posts, n (%)</td>
</tr>
<tr>
<td>Would read post</td>
</tr>
<tr>
<td>Would react to post</td>
</tr>
<tr>
<td>Would comment on post</td>
</tr>
<tr>
<td>Would share post</td>
</tr>
</tbody>
</table>
Formative research identifying the topics of interest to servers likely contributed to creating highly engaging posts in the WayToServe Plus prototype. Most servers said that they would be interested in receiving ongoing information and support for RBS methods, and many would enroll in the WayToServe Plus component in the future. Moreover, the WayToServe Plus prototype appeared to improve theoreric mediators of effective RBS training. Continuing professional development programs should be more effective when personally meaningful to learners [30,59]. Together, these findings suggest that the ongoing professional development in WayToServe Plus is likely to improve RBS practices in the upcoming RCT and when disseminated with the WayToServe RBS training.

Owners, managers, and servers felt that the professional development content fit with the atmosphere of their licensed establishments. Fit might be further enhanced by providing the information in easily digestible and relatable formats that do not require a large time commitment. Fit is an important innovation characteristic that predicts adoption [39], and continuing professional development programs may be most effective when reflecting the context and experiences of learners [30].

The formative research provided insights into the potentially effective content of a professional development component. Servers and managers wanted skills training on refusing service, handling intoxicated and difficult customers, conflict resolution, communication, drink counting, and recognizing intoxication; serving at special events; ID checking; serving laws and penalties; and prohibited conduct (eg, recreational cannabis, drinking on the job, and firearms). These represent a combination of generic as well as job-specific information and skills for alcohol service, a common combination of skills in the hospitality industry [36,37]. Servers suggested several message features that would promote engagement with the professional development content, including positive messages; relaxed, conversational tone; humor; infographics or charts; articles; videos; questions and answers; resources; reminders; interactive activities; badges or rewards; polls; quizzes; games; weekly discussion topics; tips, stories, and comments from experienced servers; and opportunities to share experiences. Sharing ideas and experiences among servers and creating learning communities where they can work collaboratively should facilitate the success of continuing professional development [39,60]. User-generated content stands out as a key feature of social media platforms and holds significant sway in shaping social norms, particularly through the process of opinion leadership [39]. Managers and servers were interested in developing professionalism, such as understanding the roles of management, building a community in the hospitality environment, and enhancing hospitality careers. Cultivating or enhancing a sense of professionalism among servers could potentially elevate their regard for customer and community safety (ie, fostering professional norms [29]), strengthen their commitment to their roles, and motivate the consistent use of RBS methods.

The findings supported the use of a social media platform to deliver the professional development content. In 2021, most adults used social media (72%), including >80% of those aged 18 to 49 years [61], for information and peer connections that can be influential [62,63]. Web-based learning is common in vocational education and continuing professional development, providing advantages in terms of low cost, time efficiency, media-rich presentations, and interactivity [31,59,60]. Our plan to deliver the professional development content on an ongoing basis should help confer mastery of RBS techniques taught initially in the single, intensive WayToServe course by providing time for servers to set goals to improve behavior, assess current performance, and receive timely feedback to make improvements [60]; however, to be effective, servers will need to be self-directed learners with sufficient motivation to engage with the post. Effectiveness and motivation may increase when coupled with in-person instruction and mentoring from managers and experienced servers [31,59,60], rather than replacing this on-the-job support.

We chose to deliver the WayToServe Plus over Facebook because (1) the WayToServe training had an existing Facebook page with approximately 20,000 followers; (2) despite some decline in its user base [64–66], Facebook still reaches a large majority of adults including more than 70% adults aged 18–49 years by one estimate in 2021 [61]; and (3) Facebook’s private group feature will control treatment presentation to prevent contamination when testing the effectiveness of WayToServe Plus. Video content appeared to be especially popular, which was not surprising given the popularity of video-dominated social media such as YouTube and TikTok [61,67]. Theoretically, visual depictions should be effective at teaching skills through observational learning [40]. To broaden the appeal of the WayToServe Plus component, some posts should be linked to relevant content posted on Instagram, YouTube, and other highly popular social media. Once disseminated, it may be most effective to deliver WayToServe Plus messaging through multiple social media platforms.

**Limitations**

The formative research and upcoming RCT evaluating the WayToServe Plus professional development component will be limited by conducting them with servers in only 3 states in the western United States, California, New Mexico, and Washington State, where the WayToServe is an approved RBS training provider. However, these states are diverse in population size, history of RBS training requirements (ie, New Mexico and Washington State have required RBS for over 2 decades, whereas California’s requirement was new in 2022), and content requirement for RBS training (eg, California requires more content for managers than New Mexico and Washington State).
The selection method using clustering of establishments in California and Washington State could introduce a design effect, but it was balanced against cost controls and project feasibility. The findings pertain to web-based professional development content, not other forms of support delivered in person or in print, but web delivery creates a high-quality, high-fidelity, engaging learning environment [68]. Uptake of WayToServe Plus will undoubtedly vary among servers and across establishments, which could diminish its effectiveness. However, the formative research suggests that many servers will engage with the professional development posts. The upcoming evaluation of WayToServe Plus will be strengthened by random selection of the licensed establishments; random assignment to experimental conditions; observational measures of refusal rates; and blending of PiP teams, establishment management, and alcohol servers.

**Conclusions**

If successful, this study has the potential to improve the effectiveness of evidence-based RBS training and reduce the negative consequences of DWI. The results will also provide evidence that personnel in regulated industries that affect public health and safety, such as hospitality, can be trained to improve compliance with state policies and regulations. Furthermore, it will show whether professional development can be effective for individuals without specialized professional education. Far from low-skilled, alcohol service requires key skills in managing emotions, communication, problem-solving, and flexibility [36] as well as learning and applying the regulations and best practices surrounding responsible alcohol service. It should be amenable to improvement through ongoing professional development between state-required retraining in RBS techniques. The market for RBS training is large; therefore, improvements in this common intervention could have a substantial impact on DWI rates. No RBS training provider currently provides ongoing professional development as extensive as is planned for WayToServe Plus, so it should be seen as a value-added component for many licensed establishments, improving its dissemination potential.

**Acknowledgments**

This study is supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R44 AA029364). The opinions are those of the authors. The funding agency had no input on the contents of this paper.

**Data Availability**

The data set generated and analyzed during the pilot study is available in the Inter-university Consortium for Political and Social Research (ICPSR) repository [69].

**Authors’ Contributions**

WGW, DB, and RS conceptualized the study, designed the methods, and secured extramural funding. WGW and DB are supervising project activities. LM is managing day-to-day study activities. All authors reviewed and approved the manuscript before submission.

**Conflicts of Interest**

WGW, DB, and LM receive a salary from Klein Buendel, Inc. DB’s spouse is an owner of Klein Buendel, Inc. WGW and DB are owners of Wedge Communications LLC, the distributor of the WayToServe web-based training. An active management plan is in place at Klein Buendel to manage this conflict of interest. RS has no conflicts of interest.

**References**


Abbreviations

- DWI: driving while intoxicated
- PIP: pseudointoxicated patron
- RBS: responsible beverage service
- RCT: randomized controlled trial
- SUS: System Usability Scale

Edited by A Mavragani; submitted 05.06.23; peer-reviewed by Y Shehawy; comments to author 13.10.23; revised version received 15.11.23; accepted 23.11.23; published 24.01.24.

Please cite as:

Woodall WG, Buller D, Saltz R, Martinez L. Professional Development to Improve Responsible Beverage Service Training: Formative Research Results and Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2024;13:e49680. DOI: 10.2196/49680. URL: https://www.researchprotocols.org/2024/1/e49680. PMID: 38265847
Novel Strategy to Assess the Neurotoxicity of Organic Solvents Such as Glycol Ethers: Protocol for Combining In Vitro and In Silico Methods With Human-Controlled Exposure Experiments

Nancy B Hopf¹,², PhD; Laura Suter-Dick²,³, PhD; Jörg Huwyler²,⁴, PhD; Myriam Borgatta¹,², PhD; Lucie Hegg¹,², MSc; David Pamies²,⁵, PhD; Hélène Paschoud¹,², MSc; Ramya Deepthi Puligilla²,⁴, MSc; Elena Reale¹,², PhD; Sophie Werne²,³, MSc; Marie-Gabrielle Zurich¹,², PhD

¹Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland
²Swiss Centre for Applied Human Toxicology (SCAHT), Basel, Switzerland
³School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Switzerland
⁴Division of Pharmaceutical Technology, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland
⁵Department of Biomedical Sciences, University of Lausanne, Lausanne, Switzerland

Corresponding Author:
Marie-Gabrielle Zurich, PhD
Department of Biomedical Sciences
University of Lausanne
Bugnon 7
Lausanne, 1005
Switzerland
Phone: 41 21 692 5542
Email: mzurich@unil.ch

Abstract

Background: Chemicals are not required to be tested systematically for their neurotoxic potency, although they may contribute to the development of several neurological diseases. The absence of systematic testing may be partially explained by the current Organisation for Economic Co-operation and Development (OECD) Test Guidelines, which rely on animal experiments that are expensive, laborious, and ethically debatable. Therefore, it is important to understand the risks to exposed workers and the general population exposed to domestic products. In this study, we propose a strategy to test the neurotoxicity of solvents using the commonly used glycol ethers as a case study.

Objective: This study aims to provide a strategy that can be used by regulatory agencies and industries to rank solvents according to their neurotoxicity and demonstrate the use of toxicokinetic modeling to predict air concentrations of solvents that are below the no observed adverse effect concentrations (NOAECs) for human neurotoxicity determined in in vitro assays.

Methods: The proposed strategy focuses on a complex 3D in vitro brain model (BrainSpheres) derived from human-induced pluripotent stem cells (hiPSCs). This model is accompanied by in vivo, in vitro, and in silico models for the blood-brain barrier (BBB) and in vitro models for liver metabolism. The data are integrated into a toxicokinetic model. Internal concentrations predicted using this toxicokinetic model are compared with the results from in vivo human-controlled exposure experiments for model validation. The toxicokinetic model is then used in reverse dosimetry to predict air concentrations, leading to brain concentrations lower than the NOAECs determined in the hiPSC-derived 3D brain model. These predictions will contribute to the protection of exposed workers and the general population with domestic exposures.

Results: The Swiss Centre for Applied Human Toxicology funded the project, commencing in January 2021. The Human Ethics Committee approval was obtained on November 16, 2022. Zebrafish experiments and in vitro methods started in February 2021, whereas recruitment of human volunteers started in 2022 after the COVID-19 pandemic–related restrictions were lifted. We anticipate that we will be able to provide a neurotoxicity testing strategy by 2026 and predicted air concentrations for 6 commonly used propylene glycol ethers based on toxicokinetic models incorporating liver metabolism, BBB leakage parameters, and brain toxicity.
Conclusions: This study will be of great interest to regulatory agencies and chemical industries needing and seeking novel solutions to develop human chemical risk assessments. It will contribute to protecting human health from the deleterious effects of environmental chemicals.

International Registered Report Identifier (IRRID): DERR1-10.2196/50300

(JMIR Res Protoc 2024;13:e50300) doi:10.2196/50300

KEYWORDS
organic solvent exposure; workers; general population; neurotoxicity; blood-brain barrier; liver toxicity; human cell cultures

Introduction

Environmental and occupational exposure to chemicals may contribute to the development of several neurological diseases [1,2]. In particular, organic solvents used in industries such as car repair, painting, furniture manufacturing, printing, and cleaning have been associated with several central nervous system (CNS) conditions. These include mild to severe toxic encephalopathy [3]; deficits in cognitive function [4-7]; and, in some cases, neurodegenerative diseases [8,9]. The diffuse neuropathological effects of acute solvent intoxication reflect neurophysiological abnormalities involving multiple brain regions. With increasingly intense or prolonged exposure, the severity of acute impairment may progress along the spectrum of delirium. Chronic high-level exposure may lead to global cognitive impairment including deficits in memory, attention, energy, and personality, which are well-described forms of dementia [10-13]. Much of the initial work on organic solvent toxicity originated in Scandinavia, where a neurobehavioral syndrome in painters leading to their early retirement was first described [14]. However, although the neurotoxicity of solvents such as toluene, trichloroethylene, and n-hexane has been recognized, the neurotoxicity of common solvents currently on the market has not been evaluated. Notably, neurotoxicity testing is only required if the chemical is deemed a pesticide; otherwise, all other chemicals are evaluated on a case-by-case basis. The only exception is if the compound structure is suspected to have nervous system targets and no data are available for read-across or when effects on the nervous system are found in single-dose (Organisation for Economic Co-operation and Development [OECD] Test Guidelines [TGs] 402, 403, 420, 423, or 425) or repeated-dose toxicity studies (TG 407 or 408). Because the nervous system effect endpoints considered as triggers (ie, modifications of wet brain weight or basic histopathology, or both) are quite insensitive, high amounts of potentially neurotoxic compounds are available on the market. The European Union Classification, Labelling, and Packaging Regulation does not include a classification for neurotoxicity. Exposure to organic solvents, especially among workers with higher exposure than the general population, can produce neurotoxic effects, depending on the internal dose. Therefore, to efficiently protect the population from possible solvent toxicity, it is important to determine the air concentrations at which neurotoxicity does not occur. To this end, we propose a strategy using a combination of in vivo (zebrafish embryo), in vitro, and in silico tools coupled with controlled in vivo human exposure experiments to assess the neurotoxicity of solvents (Figure 1).
The recognized method for the evaluation of the neurotoxic potential of chemicals, the OECD TG 424 (neurotoxicity in rodents), uses complex in vivo tests in rodents, which are laborious, expensive, difficult to apply in a standardized manner, and ethically debatable. Regulators from different agencies worldwide as well as the scientific community are becoming increasingly aware of the limitations of the current toxicity testing paradigm. Animal-based high-dose testing in typically 1 stand-alone guideline test is not always relevant for human exposure scenarios [15]. One of the most challenging aspects of this animal-centric approach is the impossibility of coping with the thousands of chemicals for which data are still lacking. Conducting animal tests is time consuming and expensive. Therefore, they cannot be carried out routinely because of the sheer number of chemicals that are currently on the market and those anticipated to enter it in the coming years [16]. In addition, there are shortcomings regarding interspecies concordance between different mammalian or rodent species as well as with respect to extrapolation from experimental animals to humans. These ambiguities in results or poor reproducibility performance call into question the relevance of such test methods for human risk assessment [16-21]. All this prompts a move away from animal testing toward a combination of in vitro and in silico approaches that address functional mechanistic endpoints [15,16,22].

Numerous in vitro models have been proposed for the evaluation of neurotoxicity in the last decades. Monolayer cultures of a single brain cell type are far from representing the human brain in terms of architecture and functionality. Given the sophistication of brain cell-to-cell interactions, some complexity is required to recapitulate human-relevant cellular processes and functions in vitro. However, a good balance must be found between this complexity and the simplicity needed to have robust and reproducible systems that can be applied for chemical screening in a high-throughput manner [23]. The 3D hiPSC–derived brain test system (BrainSpheres) we previously developed [24] will be used in this study, as it fulfills these requirements.

The blood-brain barrier (BBB) protects the brain parenchymal cells from the deleterious effects of xenobiotics. However, some chemicals are able to cross or impair the BBB [25]. The transient or permanent opening of the BBB provides xenobiotics, plasma proteins, and immunoregulatory mediators access to the CNS, where they can induce toxic effects. Therefore, we will implement a predictive model to assess the impact of solvents on BBB based on in vivo (zebrafish embryo), in vitro (human brain microvascular endothelial cells [hCMEC/D3]), and in silico models [26] to assist in the interpretation of the results obtained in in vitro neurotoxicity testing.

Glycol ethers will be used as a case study to evaluate the feasibility of our protocol. Glycol ethers form a wide family of a few dozen solvents with different physicochemical properties making them versatile and usable in a variety of industrial applications ranging from pharmaceuticals and microelectronics to domestic cleaning, personal care, and printing. The 2 main groups of glycol ethers are the E series (ie, ethylene glycol ethers [EGEs]) and the P series (ie, propylene glycol ethers [PGEs]). EGEs and PGEs show differences in their toxicological properties regarding teratogenicity, hemolysis, and testicular atrophy [23], apparently resulting from their distinct production
Healthy women and men were recruited as participants in our human study (2022-01567). Ethics committee approval was obtained from Swiss ethics committee. The toxicokinetic model will incorporate solvent air concentrations leading to brain concentrations below a toxicokinetic model in reverse dosimetry, we can predict the "pharmacokinetics," and "biokinetics") [28]. Furthermore, using a toxicokinetic model in reverse dosimetry, we can predict the solvent air concentrations leading to brain concentrations below the levels found to produce neurotoxic effects in vitro in the BrainSpheres. The toxicokinetic model will incorporate metabolism parameters derived from the in vitro liver system and passage through or toxicity to BBB. Once calibrated, the toxicokinetic model can be used to simulate chronic exposure scenarios to predict cumulative brain concentrations and used in reverse dosimetry to predict air concentrations that will not likely result in brain concentrations associated with toxicity.

Methods

Ethical Considerations

Ethics committee approval was obtained from Swiss ethics (Commission cantonale d’éthique de la recherche sur l’être humain) in 2022 for this nonclinical human study (2022-01567). Healthy women and men were recruited as participants in our study. Each participant signed a written informed consent form before inclusion in the study. The participants will be reimbursed for their time and inconvenience according to the Swiss guidelines.

Global Strategy

The choice of solvents to be included in the study will be based on the amount annually placed on the European market and the number of products registered containing known glycol ethers. The selected organic solvents will be applied to various in vitro models to determine their neurotoxicity. They must be amphiphilic to be solubilized in the cell culture media. We established the following solvent selection criteria: (1) used or produced >1 metric ton per year, (2) incorporated in numerous industrial and commercial products, and (3) water solubility. This selection process involves consulting government databases and contacting different industry sectors. We will start the project concomitantly by testing 2 solvents of the P series, propylene glycol methyl ether (PGME), for which we have already developed a toxicokinetic model, and propylene glycol butyl ether. We will test 1 additional solvent from the E series with the in vitro test system, namely, ethylene glycol methyl ether (EGME), which has been banned for use in cosmetic products in Europe [29].

The study is organized into 5 work packages (WPs). The information workflow between the WPs is shown in Figure 2.

The selected organic solvents will be applied to various in vitro models to determine their neurotoxicity. They must be amphiphilic to be solubilized in the cell culture media. We established the following solvent selection criteria: (1) used or produced >1 metric ton per year, (2) incorporated in numerous industrial and commercial products, and (3) water solubility. This selection process involves consulting government databases and contacting different industry sectors. We will start the project concomitantly by testing 2 solvents of the P series, propylene glycol methyl ether (PGME), for which we have already developed a toxicokinetic model, and propylene glycol butyl ether. We will test 1 additional solvent from the E series with the in vitro test system, namely, ethylene glycol methyl ether (EGME), which has been banned for use in cosmetic products in Europe [29].

The study is organized into 5 work packages (WPs). The information workflow between the WPs is shown in Figure 2. All results collected from the abovementioned systems will contribute to refining the toxicokinetic model we previously developed for PGME [30]. The toxicokinetic parameters of the solvent and the metabolites will then be characterized in human volunteers after exposure to PGE vapors under controlled conditions. These results will be used to calibrate and expand our toxicokinetic model [30]. The toxicokinetic model will be constructed to predict brain concentrations of selected solvents and, consequently, will include a brain compartment to predict the target organ solvent and metabolite concentrations. BBB parameters such as barrier transport, transport of the solvent once in the brain, and solvent-brain binding will be incorporated. Solvent neurotoxicity may depend on the metabolic modifications of the substances; therefore, we will incorporate the parameters for the parent compound and the metabolites found in the in vitro liver system. The data necessary to build the model will be retrieved from peer-reviewed scientific literature for tissue:blood partition coefficients (PCs) [31] following the fit-for-purpose dose-response analysis approach. The toxicokinetic model should be able to predict human brain concentration for each of the tested solvents after inhalation exposure, given the air concentration of vapors and duration of exposure. The simulated human brain concentrations will then be compared with the no observed adverse effect concentrations (NOAECs) obtained from the neurotoxicity in vitro system (BrainSpheres). In addition, the toxicokinetic model will be used to predict solvent air concentrations that are unlikely to lead to brain concentration equal to or superior to the brain NOAECs using reverse dosimetry. The specific aims of each WP are summarized in Table 1.
Figure 2. Information workflow between the work package (WP). BBB: blood-brain barrier; IVIVE: in vitro-in vivo extrapolation; NOAEC: no observed adverse effect concentration; PBTK: physiologically based toxicokinetic.
Table 1. Specific aims of the work packages (WPs).

<table>
<thead>
<tr>
<th>WP</th>
<th>Name</th>
<th>Specific aims</th>
</tr>
</thead>
</table>
| WP1     | In vitro neurotoxicity testing            | 1. Determine NOAECa for neurotoxicity of each solvent  
2. Determine the in vitro distribution kinetics of solvents  
3. Identify toxicity pathways and KEsb for solvent neurotoxicity |
| WP2     | In vivo or in vitro or in silico BBBc functionality testing | 1. Evaluate the suitability of the zebrafish embryo model to study BBB integrity and functionality  
2. Determine the impact of solvents on BBB integrity and transport in zebrafish and hCMEC/D3ed  
3. Determine the solvents permeability coefficient (Pe)  
4. Provide quantitative data on BBB permeability and tissue distribution of solvents based on computational modeling |
| WP3     | In vitro hepatic metabolism and clearance | 1. Elucidate hepatic metabolism  
2. Calculate substrate-enzymatic parameters (Vmaxe and Kmf)  
3. Detect and identify possible metabolites produced by the liver |
| WP4     | In vivo volunteer exposure                | 1. Characterize human blood absorption and urinary elimination kinetics for parent glycol ether as well as the metabolites identified in WP3  
2. Find neurotoxic and vascular injury effect biomarkers for solvent exposure |
| WP5     | In silico PBTKg modeling                  | 1. Establish and calibrate the PBTK model for various organic solvents  
2. Use reverse dosimetry to determine air concentrations below human brain toxicity concentrations |

aNOAEC: no observed adverse effect concentration.  
bKE: key event.  
cBBB: blood-brain barrier.  
dhCMEC/D3: human brain microvascular endothelial cells.  
eVmax: maximum velocity.  
fKm: Michaelis constant.  
gPBTK: physiologically based toxicokinetic.

WP1: In Vitro Neurotoxicity Testing

Testing strategies are needed to evaluate the neurotoxicity of chemicals in a more cost-effective, efficient, and ethical manner. Participating in an international effort, we developed a 3D human-induced pluripotent stem cells (hiPSC)–derived brain model containing several subtypes of neurons, astrocytes, and oligodendrocytes [24]. This system allows the cells to reach a high level of differentiation and cellular maturation, exemplified by the presence of functional synapses and compact myelin. The presence of myelin is important for this project because solvents more easily target lipid-rich structures [32]. This 3D human brain model has already proven its usefulness for neurotoxicity testing [33-36]. We hypothesized that glycol ethers are neurotoxic. Therefore, we propose to take advantage of our hiPSC-derived BrainSpheres model to study the neurotoxicity induced by uncharacterized glycol ethers present on the market, which will be compared with the neurotoxicity of well-characterized solvents known to induce human encephalopathy.

Solvents are data-poor substances. It was originally hypothesized that they exert their toxic effects largely through nonspecific physicochemical effects that modulate membrane fluidity and perturb the hydrophobic force regulating macromolecular interactions [37]. However, recent evidence supports the view that solvents interact with lipophilic areas on protein receptors [38,39]. They have also been shown to induce lipid peroxidation, leading to mitochondrial dysfunction, failure of electron transport, and energy production [40,41]. In this study, omics (eg, proteomics, metabolomics, and lipidomics) technology will be used to decipher the mechanisms of glycol ether neurotoxicity and to identify potential biomarkers of toxic effects.

Primary 3D hiPSC-derived brain cell cultures will be prepared and maintained as previously described [24]. This model contains neurons that form synapses, astrocytes, and oligodendrocytes myelinating the axons (Figure 3). Cytotoxicity will be determined by a resazurin assay after repeated exposure (7 d) to the selected glycol ethers (parent compounds and metabolites). Gene expression for cell type–specific genes, markers of synapses and myelin, and markers of cell stress will be quantified by quantitative reverse transcription polymerase chain reaction at concentrations of solvents under half-maximal effective concentration (EC50) for cytotoxicity. Immunostaining will be performed to assess the effects of solvents on synapses, myelin, and astrocyte reaction, and immunofluorescence will be quantified. NOAECs (Figure 2) will be determined for all tested endpoints, as previously shown for gene expression [42]. Brain cell cultures will also be exposed to the metabolites of PGME, propylene glycol butyl ether, and EGME and to the metabolites of the newly selected uncharacterized solvents, potentially produced by liver metabolism (WP3).
Figure 3. Brain model used in work package (WP) 1. Immunostainings of human induced pluripotent stem cell–derived 3D BrainSpheres after 8 weeks of differentiation, showing the presence of proteins specific for neurons (neurofilament heavy polypeptide [NF200]), synapses (postsynaptic density-95 protein [PSD95] and synaptophysin [SYP]), astrocytes (glial fibrillary acidic protein [GFAP]) and oligodendrocyte (proteolipid protein 1 [PLP1]). Scale bars: 40 µm.

To establish the in vitro distribution kinetics of selected solvents necessary for toxicokinetic modeling, 3D brain cell cultures and medium will be collected 3, 6, 24, and 48 hours after the first exposure and after the last exposure of the repeated treatment. The solvent and its main metabolites (if relevant) will be quantified to establish a time course of disappearance from the medium and appearance in the cells as well as to assess the potential accumulation for the entire period of exposure. The fraction bound to culture plates’ plastic will be quantified after desorption. In silico modeling of glycol ethers in vitro distribution kinetics will then be developed. This model will be able to predict the change in cell-associated concentrations of solvents in BrainSpheres with time, as previously shown for amiodarone [43].
WP2: In Vivo, In Vitro, and In Silico BBB Functionality Testing

We previously established the zebrafish as a predictive vertebrate screening model to study the systemic circulation and tissue distribution of particulate drug carriers [44,45]. At 72 hours postfertilization, zebrafish embryos have a functional CNS and, presumably, a fully functional BBB. Anatomical structures such as the vascular endothelium can be visualized using transgenic fish lines expressing fluorescent proteins (Figure 4). Defined exposure of the zebrafish can be achieved by the simple addition of solvents to the fish incubation medium within a closed container. Other advantages of the model include the possibility of studying BBB functionality under physiological conditions in vivo and the high throughput. A well-known in vitro model for the human brain endothelium, hCMEC/D3 cell line (Figure 4) showing the formation of tight junctions and the expression of most transporters and receptors of the in vivo BBB [46], cultured in a transwell system, will also be used. Furthermore, extrapolation of in silico, in vitro, and zebrafish data to higher vertebrates seems feasible [47].
Figure 4. Blood-brain barrier (BBB) models used in work package (WP) 2: zebrafish larvae (ZFL) and human brain microvasculatur endothelial cells (hCMEC/D3). ZFL (2 top panels): tracer permeability across BBB. Dorsal view of the midbrain region of the zebrafish lines Tg (kdrl:enhanced green fluorescent protein [eGFP]), which expresses eGFP (green signal) in the endothelial cell membranes. ZFL were injected with the tracer 1 kDa maleimide (red signal). Scale bars: 50 µm. hCMEC/D3 cells (lowest panel): actin filament stained with fluorescein isothiocyanate phalloidin. Scale bars: 100 µm.

Zebrafish larvae are frequently used in developmental biology or toxicological studies. However, in this study, we will use zebrafish larvae exclusively to study BBB integrity and functionality [48]. Fluorescently labeled reference compounds (Figure 4) will be intravenously injected into the Duct of Cuvier, as markers of paracellular permeability (eg, fluorescein isothiocyanate dextran 70 or fluorescently labeled liposomes), substrates of drug export transporters (eg, rhodamine-123 as P-glycoprotein substrate), or nutrient transporters (eg, fluorescently labeled transferrin as a marker for receptor-mediated transcytosis). PGME will be the first reference compound to be tested because of its high water miscibility. To precisely assess exposure, analytical methods (gas chromatography-tandem mass spectrometry [GC-MS/MS])
will be used to determine the concentrations of solvents and their metabolites in zebrafish medium, in the headspace of closed incubation vessels and tissue samples (i.e., zebrafish homogenates). Circulation, tissue distribution, and brain uptake of the reference compounds will be monitored by confocal laser scanning microscopy (live imaging of anesthetized fish embryos for up to 24 hours). The concentration-dependent toxicity of solvents or their metabolites will be monitored based on the viability and malformations of embryos. The integrity of vasculature will be visualized in transgenic zebrafish kdrl: enhanced green fluorescent protein embryos. The metabolic capacity of the zebrafish will be determined by quantifying potential metabolites (determined in WP3; Figure 2) in zebrafish tissue homogenates. The concentration-dependent toxicity to BBB and the coefficient of permeation of solvents will additionally be evaluated in the hCMEC/D3 cell line cultured in a transwell system.

Finally, quantitative estimates of passive cellular uptake and BBB permeability of solvents and their metabolites will be provided based on computational modeling using physicochemical molecular descriptors according to the methods we previously established [26,49]. These methods provide very high throughput, allowing the screening of web-based chemical libraries.

**WP3: In Vitro Hepatic Metabolism and Clearance**

Because the liver is the main organ responsible for metabolism and a large contributor to compound clearance, we will implement a system suitable for predicting the hepatic metabolism of solvents. In recent years, 3D liver cell models have been proposed as an alternative to less physiological 2D cell monolayers, and their applications have progressed substantially [50]. They are widely used for the assessment of hepatotoxicity [51-55]. An advantage of spheroids is that they overcome the limitation of rapid decline of drug-metabolizing enzyme activities in primary human hepatocyte suspension culture and cell lysates [56], such as microsomes and liver S9 fractions. In this study, we will use 3D liver cultures of the well-characterized human HepaRG cell line (Figure 5), which represent a promising model to evaluate hepatotoxicity and hepatic metabolism [57].
Figure 5. Liver model used in work package (WP) 3. Bright field and immunostainings of liver 3D HepaRG cultures showing the presence of albumin and cytochrome P450 3A4 (CYP3A4). Bars: 100 µm.

Determining the appropriate experimental test system (e.g., cell plate, incubation time, and exposure concentration) will be an essential part of the development of the 3D model. Moreover, analytical methods (GC-MS/MS and liquid chromatography-tandem mass spectrometry) to detect and quantify the solvents and the metabolites formed must be developed to calculate hepatic metabolism and clearance. Then, proof of metabolic competence and maintenance of the 3D HepaRG cells will be carried out by assessing the metabolism of known P450 substrates. The presence and secretion of albumin as a specific hepatocyte marker will be assessed using immunostaining and enzyme-linked immunosorbent assay. Solvent- and metabolite-induced cytotoxicity will be assessed after 48 hours and 7 days (repeated) of exposure to determine the nontoxic concentration range for subsequent experiments. The metabolic abilities of the 3D HepaRG model and 3D
primary human hepatocytes will be compared. Furthermore, the clearance data obtained from the 3D HepaRG model will be compared with the short-term clearance measured in the human liver cell lysate (S9 fractions). In addition, Michaelis-Menten-Kinetic parameters ($V_{\text{max}}$ and $K_m$) for the formation of the metabolites will be derived using the S9 fractions. These data will be used to build a physiologically based toxicokinetic (PBTK) model (Figure 2).

**WP4: In Vivo Volunteer Exposure**

Human biomonitoring refers to monitoring exposure-related health risks by analyzing biological samples, usually blood and urine samples [58]. The biomonitoring limit values (BMLVs) are set to protect human populations against the potential toxic effects of chemical substances. These limit values account for all routes through which a chemical can enter the body. These are most often the inhalation and skin routes in occupational and environmental settings. Kinetic studies that provide absorption, biotransformation, and elimination rates as well as the absorption and elimination half-lives of the parent compound and its metabolites are necessary to set BMLVs. The apparent urinary elimination half-lives of the parent compound and its metabolites will later be used to develop a biomonitoring method. Sample collection time is crucial and is determined by the apparent elimination half-life of the chemical. Blood concentrations will be used to calibrate the air:blood PC for the toxicokinetic models.

We will recruit 4 participants for 2 of the selected solvents. All participants must meet the following criteria: they should be healthy individuals who do not smoke or use contraceptive hormones, do not consume alcohol, be aged between 18 and 65 years, have normal red blood cells and hemoglobin concentrations, maintain a BMI between 18 and 25, and should not be working with glycol ethers. Pregnant and breastfeeding women will be excluded from this study. Participants will be recruited using flyers and announcements distributed at the teaching hospital, university websites, and bulletin boards. All participants will sign a written informed consent form before being included in the study.

The participants will be exposed to a single glycol ether for 4 hours under controlled conditions in an exposure chamber (12 m$^3$). PGE concentrations will be set at or below the Swiss occupational exposure level (OEL) if one exists. In the absence of an OEL, we will rely on existing OELs for other propylene glycols. The parent compound (free and conjugated) and the oxidative metabolites (free and conjugated) of the selected glycol ethers will be monitored in blood, urine, and exhaled air samples. These are noninvasive methods used for human participants, and the results will be used in WP5 to estimate brain concentrations. All compounds will be quantified using capillary gas (parent compound in blood, urine, and exhaled air) or liquid (metabolites in blood and urine) chromatograms with tandem mass spectroscopy detection.

**WP5: In Silico PBTK Modeling**

PBTK models can be used to estimate human brain concentrations. The risk of neurotoxic effects can be estimated by comparing the predicted solvent-brain concentrations with the NOAEC obtained from the in vitro models. Mathematical models such as PBTK models can be used to predict the ADME of a chemical and its metabolites. In these PBTK models, the body is represented by 1 or more compartments. Each compartment represents 1 or more tissues that are kinetically homogeneous, that is, that have similar perfusion rates and an assumed similar substance solubility. PBTK models are described by a set of parameters that define the compartments and a set of mass balance differential equations for each compartment.

A previously developed toxicokinetic model for PGME with metabolism that is assumed to follow Michaelis-Menten kinetics calibrated for different age groups serves as the basis for our development [30,31,59]. We aim to modify this previously developed toxicokinetic model to include a separate compartment for the brain using BBB flux rates obtained from in vivo, in vitro, and in silico models (WP2). In addition, we will implement PC obtained from empirical human experiments (WP4) and metabolic parameters assessed in a hepatocyte assay (WP3). The toxicokinetic models will be able to simulate not only acute but also chronic exposures; therefore, both short-term and long-term exposures can be explored in silico. We will develop the toxicokinetic model into a physiologically based pharmacokinetic model based on the existing inhalation-only toxicokinetic model originally developed for PGME [40] and build it in the Berkeley Madonna software or equivalent. We will model the brain as a single compartment with direct contact with the blood flow and where organic solvent uptake will be assumed to be diffusion limited, which is in line with other physiologically based pharmacokinetic models [60]. Values for physiological parameters (volume of vascular brain, as fraction of brain volume [FVbh], volume of extravascular brain, as fraction of brain volume [FVbvh], volume fraction of brain tissue [FVB; as percent of body weight], BBB surface [Sh] in cm$^2$, fraction of cardiac output in brain at rest [BFbrainrest]/cardiac output in brain at light work [BFbrain]) required to build the TK model are from the scientific literature. Depending on the substance, values of chemical-specific parameters such as the pulmonary retention (Rpulm), central:air PC (Pca), blood:air PC (Pba), and brain tissue:vascular brain PC (Pevb vb) are either taken from the literature or estimated in silico. Since the partitioning of organic compounds between human tissue homogenate and blood is a function of water and lipid content of tissues and the n-octanol:water PC (Kow), PCs are estimated in silico based on LogKow. Kinetic coefficients needed for each organic solvent included in this study will be found in WP3 for liver metabolism (Michaelis-Menten parameters [$V_{\text{max}}$ and $K_m$]), WP2 for BBB uptake (BBB permeability-surface area product [PS]). The fraction unbound in blood (Fu_blood) will be estimated based on the fraction unbound in plasma (Fu_plasma) and the blood-to-plasma ratio (Rb), and the fraction unbound in brain (Fu_brain) will be considered when modeling each solvent as only the free fraction is able to distribute to different tissues and is biologically active. The model will be calibrated by comparing the predicted and actual urinary organic solvent concentrations obtained from the controlled human experiments (WP4). Both the free and total
organic solvent concentrations (free+conjugated) will be obtained for calibration.

**Results**

With this project, we expect to provide a strategy to rank uncharacterized solvents and their potential liver-formed metabolites, according to their potential neurotoxicity, and in comparison with the banned EGME. More importantly, a series of PBTK simulations will be conducted to predict occupational exposure, assuming 8 hours of exposure per day, 5 days per week, physical activity for 12 hours per day, and rest for the remaining 12 hours. We will use the PBTK model in reverse dosimetry to estimate air concentrations that do not produce brain concentrations determined as neurotoxic in the hiPSC-derived 3D brain model. We will recommend that authorities setting occupational exposure and public health limits consult these values. Keeping the exposure below the brain effect level should ultimately increase the protection of exposed workers and the general population with domestic exposures.

We also anticipate gaining insights into the mechanisms of action of solvents of the glycol ether family. We will elucidate the possible toxic endpoints in the brain, liver, and zebrafish models. Furthermore, we will be able to establish how toxicity is related to the compounds’ lipophilicity and metabolites.

**Discussion**

Overall, our strategy combining multiple, fit-for-purpose 3D advanced cell culture systems; zebrafish larvae; biomarker analysis; human ADME experiments; and in silico prediction is expected to contribute to the improvement of human risk assessment. Although we identified some risks we could encounter during the project, we are confident that our already determined mitigation measures will be able to overcome potential pitfalls.

Determination of the passage of solvent through the BBB may be challenging; hence, we are applying 3 different complementary methods: in vivo zebrafish larvae, in vitro human cells (hCMEC/D3), and in silico models. We are also considering and assessing the effects of the hepatic metabolites of the solvents on human BBB cells. With this experimental strategy, issues regarding the potential direct effect of solvents on cell membranes, the relatively low miscibility of solvents with water, and the physiological differences between zebrafish and humans (eg, metabolism and route of expected) should be overcome.

We have extensive experience in recruiting human volunteers for controlled human exposure sessions in the exposure chamber. Sometimes, recruitment takes longer than anticipated, and if that is the case, we will extend the timeline to not compromise the size of the study. New analytical chemical methods will need to be determined, which is time consuming. However, we will use a laboratory with extensive experience in analyzing PGME in urine and blood samples. This will also have to be accommodated with a delay in the timeline.

Future developments, not included in this study, are a strategy extended to include developmental neurotoxicity by determining other endpoints, such as proliferation and neurite outgrowth, after exposure of BrainSpheres to solvents at earlier developmental stages and by adding an in vitro test system to take into account the passage of solvents through the placental barrier [61]. We might also consider combining zebrafish embryo behavioral assays (eg, spontaneous tail coiling) with the BrainSpheres model as readouts for developmental neurotoxicity. Finally, the PBTK model could be adapted to determine solvent air concentrations that are unlikely to cause neurotoxic effects in fetuses or pregnant women.

**Acknowledgments**

This work was partially funded by the Swiss Centre for Human Applied Toxicology (SCAHT), Basel, Switzerland.

**Data Availability**

Data will be available from the corresponding author on reasonable request.

**Authors' Contributions**

NBH, LSD, JH, and MGZ conceptualized the paper. NBH, LSD, JH, and MGZ wrote the original draft. All authors reviewed and edited the paper. LH, DP, HP, RDP, and SW visualized the paper. NBH, LSD, JH, and MGZ administered the project. NBH, LSD, JH, and MGZ acquired the funding.

**Conflicts of Interest**

None declared.

**References**


53. Riede J, Wollmann BM, Molden E, Ingelman-Sundberg M. Primary human hepatocyte spheroids as an in vitro tool for investigating drug compounds with low clearance. Drug Metab Dispos 2021 Jun 01;DMD-AR-2020-000340 [**FREE Full text**] [doi: 10.1124/dmd.120.000340] [Medline: 34074732]


60. Hopf et al. JMIR Res Protoc 2024 | vol. 13 | e50300 | p.45https://www.researchprotocols.org/2024/1/e50300

**Abbreviations**

ADME: absorption, distribution, metabolism, and excretion
BBB: blood-brain barrier
BFbrain: cardiac output in brain at light work
BFbrainrest: cardiac output in brain at rest
BMLV: biomonitoring limit value
CNS: central nervous system
EGE: ethylene glycol ether
EGME: ethylene glycol methyl ether
Fu_blood: fraction unbound in blood
Fu_brain: fraction unbound in brain
Fu_plasma: fraction unbound in plasma
FVvb: volume of vascular brain, as fraction of brain volume
FVevb: volume of extravascular brain, as fraction of brain volume
GC-MS/MS: gas chromatography-tandem mass spectrometry
hiPSC: human induced pluripotent stem cell
NOAEC: no observed adverse effect concentration

https://www.researchprotocols.org/2024/1/e50300
OECD: Organisation for Economic Co-operation and Development
OEL: occupational exposure level
Pha: blood:air partition coefficient
PBTK: physiologically based toxicokinetic
PC: partition coefficient
Pca: central:air partition coefficient
Pevb_vb: brain tissue:vascular brain partition coefficient
PGE: propylene glycol ether
PGME: propylene glycol methyl ether
PS: blood-brain barrier permeability-surface area product
Rb: blood-to-plasma ratio
Rpulm: pulmonary retention
Sh: blood-brain barrier surface
TG: Test Guideline
Vmax: maximum velocity
WP: work package
Implementation of a Primary Prevention Program for Posttraumatic Stress Disorder in a Cohort of Professional Soldiers (PREPAR): Protocol for a Randomized Controlled Trial

Emmanuelle Le Barbenchon1*, PhD; Marion Trousselard2*, MD, PhD; Sonia Pellissier1, PhD; Mathilde Moisseron-Baudé4,5,6, PhD; Philippine Chachignon1, MAST; Pierre Bouny7, PhD; Emma Touré Cuq7, PhD; Sandrine Jacob3; Cécile Vigier2; Maud Hidalgo8, PhD; Damien Claverie2*, MD, PhD; Anais M Duffaud2*, PhD

1Laboratoire de Psychologie Sociale, Aix Marseille Université, Aix-en-Provence, France
2French Armed Forces Biomedical Research Institute, Brétigny-sur-Orge, France
3Univ Savoie Mont Blanc, Univ Grenoble Alpes, LIP/PC2S, Grenoble, France
4School of Practicing Psychologists, Paris, France
5VCR Team, School of Practicing Psychologists, Catholic Institute of Paris - Religion, Culture and Society Reception Team, Paris, France
6Center for Research on Work and Development, Conservatoire National des Arts et Métiers, EA4132, Paris, France
7URGOTECH, Paris, France
8Centre d’Épidémiologie et de Santé Publique des Armées, Marseille, France
*these authors contributed equally

Corresponding Author:
Anais M Duffaud, PhD
French Armed Forces Biomedical Research Institute
1, Place Général Valérie André
Brétigny-sur-Orge, 91223
France
Phone: 33 0618942117
Email: anaisduf2aud1@gmail.com

Abstract

Background: Posttraumatic stress disorder (PTSD) is a psychiatric disorder that can manifest after a traumatic event where the individual perceives a threat to his or her life or that of others. Its estimated prevalence in the European population is 0.7% to 1.9%. According to the “dose-response” model, individuals who are most exposed to traumatic events are most at risk of developing PTSD. Hence, it is unsurprising that studies have observed a higher prevalence among the military population, ranging from 10% to 18%, or even up to 45%. This project’s overall goal is to evaluate the primary prevention actions that can strengthen the resilience of at-risk professionals, notably military personnel, in the short term, with the medium- to long-term aim of preventing the occurrence of PTSD and improving the patient’s prognosis.

Objective: This study’s objectives are (1) to design a primary prevention program for PTSD, tailored to the studied military population and compatible with operational constraints; and (2) to implement and validate the Primary Prevention of Posttraumatic Stress Disorder in Military Professionals (PREPARE) program in the short term with operational personnel belonging to the French Mountain Infantry Brigade.

Methods: This is a single-center, prospective, randomized, parallel-group controlled cohort study. The cohort is divided into 2 groups: the nonintervention group receives no training, and the intervention group follows a dedicated prevention program (structured into 8 workshops and 2 debriefing and practice reinforcement workshops). Each participant is evaluated 4 times (at inclusion, +4 months, +6 months, and +12 months). During each visit, participants complete several psychosocial questionnaires (which take 15-80 minutes to complete). Samples (a 30-mL blood sample and three 5-mL saliva samples) are collected on 3 occasions: at inclusion, +4 months, and +12 months. Emotional reactivity (electrocardiogram and electrodermal activity) is measured before, during, and after the classic and the emotional Stroop task.

Results: The project is currently ongoing, and results are expected to be published by the end of 2024.

Conclusions: The study adopts an integrative approach to the processes that play a role in the risk of developing PTSD. Our biopsychosocial perspective makes it possible to target levers related to factors specific to the individual and socio-professional activities.
factors. The following dimensions are addressed: (1) biophysiology (by studying markers of the neurobiological stress response, wear and tear, and vulnerability phenomena and reinforcing the flexibility of the autonomic nervous system), (2) psychology (by facilitating and measuring the development of flexible coping strategies to deal with stress and evaluating the moderating role of the individual’s sense of duty in the development of PTSD), and (3) social (by facilitating community strategies aimed at reducing stigmatization and supporting the use of care by professionals in difficulty, in the institutional context).

**Trial Registration:** ClinicalTrials.gov NCT05094531; https://clinicaltrials.gov/study/NCT05094531

**International Registered Report Identifier (IRRID):** DERR1-10.2196/47175

**KEYWORDS**
posttraumatic stress disorder, military, primary prevention, biopsychosocial, resilience, coping, stigma, biophysiology; PTSD; implementation; soldier; veterans; prevention program

**Introduction**

**Background and Rationale**

**Overview**
Posttraumatic stress disorder (PTSD) is a psychiatric disorder that manifests following the experience of a traumatic event (TE) where the individual has perceived a threat to his or her life or that of others [1]. In France, it is considered to be the third most widespread psychiatric disorder, after major depression and specific phobias. Its prevalence in the European population is estimated to be between 0.7% and 1.9% [2,3]. According to the “dose-response” model, individuals who are most exposed to TEs are most at risk of developing PTSD [3-5]. At-risk occupations, such as the military, law enforcement, and first responders, carry an inherent risk of experiencing trauma. Hence, it is unsurprising to observe a high prevalence of PTSD in military populations, ranging from 10% to 18% or reaching up to 45% [5-9], depending on the study. Although cumulative exposure seems to be an important determinant in the general population [10], the literature does not establish a clear link among at-risk professionals. Nevertheless, it is reasonable to think that cumulative exposure is detrimental, and may even lead to fragilization, which would be consistent with the intensity of PTSD symptoms [11,12]. Repeated exposure triggers a reaction that builds on the consequences of previous exposures, increasing the complexity of the link between cumulative exposure and PTSD development, via multiple pathways.

**PTSD Prevention**
New medical research has influenced public policy, and the promotion of good health has become a key priority where policies seek to encourage certain behaviors that protect, improve, or restore the health of individuals, groups, or the entire population, while a prevention framework is used when the goal is to limit or prevent the development of a specific disease.

In the 1980s, evidence-based medicine emerged. The approach is founded on the principle of structuring prevention and public health decisions based on scientific evidence (eg, epidemiological studies and systematic clinical studies). A given health problem can be addressed by preventive interventions at primary, secondary, or tertiary levels. The objective of primary prevention is to intervene before the problem occurs; actions target the determinants of the problem by reducing risk factors or by promoting or reinforcing protective factors. In the context of PTSD, primary prevention can intervene at the level of preventing the trauma or at the level of PTSD prevention after the TEs. For at-risk professions in particular, PTSD prevention is the primary target, as the risk of exposure to trauma is integral to the profession. Secondary prevention seeks to define interventions once the problem is identified; it consists of early detection and referral for treatment (ie, after exposure to a TE). Finally, tertiary prevention seeks to manage the problem and prevent relapse. If necessary, the goal is to treat or limit any aggravation of the problem, notably any psychosocial consequences or comorbidities associated with PTSD. At all 3 levels, prevention must respect individual freedom [13].

**Primary Prevention of PTSD**
Effective primary prevention relies upon a sound theoretical understanding of the processes and determinants of the problem, knowledge of actions that have been scientifically validated in other populations, and clinical expertise in the field, which reflects the characteristics (eg, sociocultural) of the target population. This “selective prevention” approach seeks to target exposed participants.

In the domain of PTSD, primary interventions are rare, as such actions aim to prevent the impact of TEs before they occur. Nevertheless, this literature [14] gives us an insight into the potential effects of cumulative exposure to TEs during the career of military professionals. In this context, primary prevention interventions target several determinants that contribute to developing resilience following the TE. The approach underlines the need to characterize the factors that support individual and collective resilience processes, tailored to the characteristics of the population of at-risk professionals. This initial step is a prerequisite to the definition of relevant targets and the development of an appropriate prevention intervention for the population.

The proposed program (which we call Primary Prevention of Posttraumatic Stress Disorder in Military Professionals [PREPAR]) is part of a comprehensive prevention approach that synergistically links physiological, psychological, and social determinants. The approach stems from the etiological model proposed by Jones and Barlow [15]. This model is a comprehensive framework used in clinical psychology to understand the development and maintenance of psychological problems.
disorders, particularly PTSD. It integrates various factors that contribute to the onset and perpetuation of these disorders. Several etiological factors, including biological, cognitive, and behavioral components, are implemented in this model. The model also takes into account predisposing factors and moderating variables. It emphasizes a holistic approach to understanding mental disorders, taking into account the complexity of the human experience.

The Physiological Dimension

PTSD is described as a failure of emotional extinction which develops following exposure to intense fear. Exposure to a stressful event leads to a neurobiological stress response, resulting in the activation of the sympatho-adrenergic neurovegetative system and the corticotrophic neuroendocrine axis. Although this dual physiological response is effective in the short term, it comes with a biological cost; regulatory mechanisms aim to compensate for this loss by supporting poststress recovery. A repeated inability to recover and extinguish the stress response in the long term (repeated exposures over a short time period) creates a so-called “allostatic load” [16], which progressively limits the flexibility of the central and peripheral nervous system. The peripheral nervous system is of particular interest, due to its role as a mediator of the allostatic load. The parasympathetic branch plays a role in the emotional extinction and correction of the load, making it a key vulnerability factor for health when it is insufficiently effective [17]. Certain professions (frontline responders and military personnel) are at risk of developing a significant allostatic load, due to their repeated exposure to TEs (inherent in the nature of their work). On top of this, personnel must adapt, on a daily basis, to various environmental demands: physiological (sleep debt, altered sleeping patterns, etc), physical (hypoxia, etc), and cognitive and emotional (traumatic exposure, etc), among others. All of these factors test the flexibility of the individual’s physiological systems, particularly the autonomic (ie, parasympathetic) nervous system, which, consequently, appears to be a key target for prevention measures.

The Psychological Dimension

Military personnel participate in multiple missions in conflict zones, and this is likely to alter the flexibility of their executive and emotional regulation functions. Although training programs aim to develop automated responses to well-known mission scenarios, they do not focus on developing the change in viewpoint that contributes to the resilience process. Moreover, preclinical data show that both acute and repeated stress are likely to reduce cognitive flexibility [18,19]. A slight reduction in cognitive flexibility likely reduces the flexibility of emotional regulation strategies following a TE, thus increasing the impact of exposure. This hypothesis draws upon the notion of “coping flexibility.” The latter concept represents the individual’s ability to evaluate the effectiveness of his or her strategies for coping with stress in a specific situation and then adopting alternative strategies, if necessary [20]. The evaluation of training in the diversification of emotional regulation strategies in military populations has found a reduced risk of PTSD after exposure to TEs. Furthermore, a relationship has been identified between cognitive flexibility and self-compassion [21]. The data also suggest that increasing self-compassion contributes to increasing cognitive flexibility [22].

Finally, a clear sense of duty has been found to counterbalance perceived constraints associated with the mission [23,24]. The literature reports that the meaning attributed to the work environment can, to a significant degree, compensate for the perception and impact of occupational stressors [25].

The Social Dimension

The social dimension targets normative pressure and stigmatization. These determinants are specific to the institutional context of military personnel and may represent risk factors. Thus, normative group pressure to follow codes that encourage the nonexpression of emotions [26] can be a risk factor, as military personnel can be reluctant to speak about their psychological and somatic symptoms, which delays treatment. A fear of stigmatization by the institution and peers is another potential barrier to care. Overall, these social factors are obstacles to both individual and collective positive health behaviors.

Research Hypotheses

Overview

The goal of this project is to evaluate a multidimensional biopsychosocial primary prevention intervention for PTSD aimed at strengthening the resilience of at-risk professionals, namely military personnel, in order to prevent the occurrence of PTSD or reduce its severity.

The objectives of the project are as follows: (1) to design a primary prevention program for PTSD specific to the studied military population and compatible with operational constraints; (2) to implement or validate the program with operational personnel belonging to the French Mountain Infantry Brigade (Brigade d’Infanterie de Montagne); and (3) to understand PTSD and its prevention from 3 perspectives: biophysiological (by studying key markers of the neurobiological stress response, strain and vulnerability and increasing the flexibility of the autonomic nervous system); psychological (by facilitating and measuring the development of flexible strategies to cope with stress and evaluating the moderating role of the meaning of the mission in the development of PTSD); and social (by facilitating community strategies aimed at reducing stigmatization and helping professionals in difficulty to access care in the institutional context).

Our biopsychosocial approach adopts an integrative understanding of the processes at play in the risk of developing PTSD. This perspective makes it possible to target levers related to factors specific to the individual (at physiological and psychological levels), contextual and social factors (related to the working environment).

The Physiological Dimension

The change in parasympathetic vagal flexibility is an early and silent sign of physiological deterioration. It can be identified by noninvasive measurements (resting or tonic heart rate variability, and activation or phasic heart rate variability) and modulated by exercises that target vagal activity (cardiac...
coherence techniques), which can be easily integrated into a busy professional agenda. A preliminary feasibility study among a group of firefighters (SDIS 73) demonstrated very good acceptance of this technique by the selected professionals and good adoption in daily life [24].

The Psychological Dimension

Two factors are targeted in the psychological dimension: coping flexibility and the sense of duty. These 2 determinants have rarely been targeted in studies of the prevention of occupational PTSD. The focus on coping flexibility and self-compassion seeks to improve emotional regulation, and the focus on the sense of duty seeks to target the silent determinants of occupational PTSD, with the overall aim of identifying an effective primary prevention intervention.

The Social Dimension

The targeted determinants are the beliefs and socio-normative processes that play a role in PTSD in a military population (self-stigmatization, stigmatization by others, normative pressure, and group cohesion). The intervention aims to establish a dedicated space to normalize the discourse, with the overall aim of discussing PTSD and making it visible, along with the signs of emerging psychological and somatic injuries. Specific group facilitation techniques are used, notably social modeling, which aims to reinforce feelings of self-efficacy that support the public expression of signs of injury and to support the search for appropriate care, when necessary [27].

Objectives of the Research

The main objective is to determine the effectiveness of a biopsychosocial program targeting the resilience (developing stable resources to adapt to occupational demands, in order to cope with TEs) of military personnel in the context of PTSD prevention.

The secondary objective is to better understand interindividual variability regarding the program’s impact. We will study levels of vulnerability at inclusion and their impact on the program’s effectiveness. We will also track postprogram changes in resilience (at 6 and 12 months). In addition, we will evaluate (1) the impact of the program on activation and psychobiological deterioration biomarkers and (2) the psychosocial determinants at 4 months (the end of the program) and 12 months (to evaluate persistence). In addition, we will measure adherence to the intervention at group and individual levels, through an evaluation at the end of the program. Finally, we will examine user satisfaction with the URGOfeel sensor and its application during workshops.

Methods

Ethical Considerations

Prior to their participation in the study, all participants will receive 2 separate notifications: one detailing the primary study and the other focusing on the genetic part. It is essential to emphasize that participation in the genetic segment is entirely optional, allowing individuals to engage solely in the main study if they prefer. Both of these notifications provide comprehensive information about the study’s objectives, limitations, and legal requirements, especially in terms of privacy and confidentiality. Additionally, they clearly outline the potential benefits and risks associated with participation.

All collected data will undergo rigorous anonymization and processing in strict accordance with the MR001 reference methodology, in compliance with the regulations of the French Data Protection Authority (CNIL).

Our study maintains a steadfast commitment to ethical standards, in alignment with the principles set forth in the 1964 Helsinki declaration and its subsequent amendments. Furthermore, the ethics committee of Ile de France 8 formally approved the research protocol on January 12, 2021 (reference 20.12.10.58611). It is important to note that participants will not receive any form of compensation for their participation in this research.

Recruitment

The study will be conducted with members of the 27th Mountain Infantry Brigade, located in the Auvergne-Rhône-Alpes region of France. Participants will be recruited from companies proposed by the Brigade’s line manager, depending on the ability to operationalize and implement measurement sessions, and the intervention.

Randomization

Individual randomization is not feasible for two reasons: (1) the logistics involved in deploying the training program and the availability of participants; and (2) the need to avoid a contagion effect between members of the Brigade who benefit from the program and those who do not.

Eligibility Criteria

After obtaining informed, written consent, the following inclusion criteria will be verified: affiliated to a social security scheme; a male adult; a member of a combat unit with external operations (OPEX) capability; a soldier with a current contract with the Brigade, lasting a minimum of 12 months; a soldier who is able to attend all measurement sessions and workshops, according to the schedule defined upstream; a soldier who is not a member of the Groupements Commando Montagne (to ensure homogeneity); and, finally, nonparticipation in one of the studies included in phase 1 of the PREPAR project. The latter social psychology study aimed to understand the day-to-day experience of frontline professionals and was approved by the ethics committee of the University of Aix-Marseille (2019-12-12-001). Noninclusion criteria are as follows: being female; receiving treatment for a chronic disorder (daily medication for at least 1 month); participation in an external operation planned within 12 months; and being an adult ward of court.

Participant Timeline

Figure 1 presents the timeline for participants, and Table 1 presents a synopsis of how the study will unfold.
The first step is a briefing session with 2 of the Brigade’s companies. Each individual will receive 2 information letters (1 regarding participation in this study, and 1 regarding participation in genetic studies), in order to allow time to reflect before the inclusion visit.

**Visit 1: Inclusion**

Inclusion will take place over a 3-week period. This time is needed to be able to disseminate information to the groups concerned and collect data. Once informed consent has been obtained, and eligibility criteria have been verified by one of the investigators, the following samples will be collected: (1) a 30-mL blood sample; (2) three 5-mL saliva samples (the participant will be asked to collect a sample on waking, getting up, and 30 minutes later [28]); (3) noninvasive measurements of cardiac variability and electrodermal conductance (using patches applied to the skin) during Stroop tasks, estimated to take 20 minutes; and (4) a psycho-cognitive evaluation (standardized and validated questionnaires). The latter will make it possible to evaluate the psychological and somatic symptoms (Patient Health Questionnaire [PHQ-15] [29]; psychosocial factors; Perceived Stress Scale [PSS] [30]; coping flexibility [20]; inner correspondence and peaceful harmony [ICPH] [31]; stigma and barriers to care [32]; Multidimensional Scale of Perceived Social Support [MSPSS] [33]; the Siebold vertical and horizontal cohesion questionnaire [34]; and the presence of any psychopathologies such as anxiety and depression: Hospital Anxiety and Depression Scale [HADS] [35,36], Posttraumatic Checklist-5 [PCL-5] [37], unsure, and Burnout Measure Short Version [BMS] [38,39]). The time required to fill out the questionnaires is estimated to be 45-60 minutes (184 items in total).

### Table 1. Summary of the stages of the study.

<table>
<thead>
<tr>
<th></th>
<th>Visit 1 inclusion</th>
<th>+4 months</th>
<th>+6 months</th>
<th>+12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information – consent</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociodemographic data</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Self-administered questionnaires</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Blood and saliva collection (physio-biology and genetic)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Emotional reactivity</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>“Debriefing and anchoring of the practice” workshop</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**The Prevention Program: Workshops**

Members of the intervention group will participate in 8 workshops. Each lasts 2 hours and is divided into three parts: (1) welcome participants with time allocated for inclusion (time approximately 20 minutes); (2) a presentation of the pedagogical objectives of the session, and its practical application; and (3) group wind-up (approximately 20 minutes). Please see Multimedia Appendix 1 for a description of the workshop program. At the end of the study, members of the nonintervention group, who wish to do so, will be able to benefit from the same prevention program during a later session, which will be scheduled as a function of operational constraints.

**Visit +4 Months: Follow-up**

A second round of data collection will take place at the end of the prevention program (approximately 4 months after inclusion). Each participant will be asked to allocate half a day to the collection of biopsychological measurements.

**Visit +6 Months: Follow-up**

Two months after the end of the program, during the first debriefing workshop, participants will be asked to complete a set of self-administered questionnaires (the same set as the one completed during the inclusion visit a part for the sociodemographic information).

**Visit +12 Months: End of Study**

Eight months after the end of the prevention program, each participant will take part in a final measurement session (lasting 2-3 hours). They will also be asked to allocate another half a day to the collection of biopsychological measurements.
Measurements recorded at +4 and +12 months are identical to those performed at inclusion, except for sociodemographic data, and the sample needed to assess vulnerability (genetic polymorphism).

Thus, all participants are required to attend 4 measurement sessions: visit 1, inclusion; +4 months; +6 months; and +12 months. During each of these visits, they will be asked to complete a self-administered (using a tablet) questionnaire to collect psychosocial data. In addition, at inclusion, +4 months, and +12 months, (1) biological samples will be collected for each participant (blood and saliva) and (2) activation of the autonomic nervous system (heart rate variability and electrodermal activity) will be measured at rest, during emotional activation (the classic, then the emotional Stroop task), and at recovery.

Outcomes

Primary End Point
There is no consensus in the literature regarding the definition of resilience, and there are references to several dimensions. Consequently, measuring the effect of the intervention on resilience requires the use of several indicators. A simple approach would be to measure PTSD symptomatology reported by participants, with the expectation that the intervention will result in a decrease in symptoms. Although this approach would be relevant for long-term measures [40], using this criterion in a short-term study, such as ours, would be restrictive and fail to support the processes that are assumed to be activated during the intervention [41]. The latter consists, in part, of removing barriers to seeking care when necessary, in order to support the building of resilience as a process, and not simply as a state. Our short-term approach is also justified by the difficulty of conducting studies in a military environment over the long term (personnel are transferred every 3 years) and attrition among professional soldiers.

The proposed composite criterion groups indicators of resilience that are described as protective factors for PTSD and are recognized to be sensitive to interventions. The first is the participant’s emotional state. It is known to be highly impacted in PTSD, where there is an increase in negative emotions. At the same time, it is a marker of resilience in cases where emotions become positive following trauma. We will therefore use the Positive and Negative Affect Scale (PANAS) score [42], which measures the intensity of positive and negative emotional states. Our second indicator measures self-compassion, which refers to a general tendency to be kind to oneself, despite knowing one’s failures and successes. Self-compassion is very directly linked to the process of self-acceptance and is a predictive criterion for both the development of PTSD (when it is deficient), and increased resilience (when stimulated by a psychotherapeutic intervention). Self-compassion also contributes to cognitive flexibility. Finally, the third indicator is the hardness score. This measure is classically used in the literature to study, from a psychological point of view, an individual’s ability to remain in good health under stressful conditions. It should be noted that this measure should not, by itself, be considered a sufficient criterion for resilience in our protocol, given that any modification is observed over the long term, based on a quasi-dispositional approach.

Consequently, the hardness score will be combined with mood and self-compassion criteria. The composite criterion will be evaluated at 4 months (Figure 1). A change in resilience will be defined as favorable if at least two of the following three criteria are met: (1) a 20% improvement in the PANAS score; (2) given the military context, and the nonspecific nature of our intervention, a 20% improvement in scores on the Self-Compassion Scale; or (3) a 5% change in hardness, measured using the Dispositional Resilience Scale (DRS-15) [43]. If these criteria are not met, any change will be considered to be unfavorable. The choice of thresholds for variables making up the composite criterion is based on data from research conducted by the project’s teams.

Emotional State
The data based on the STEP study, carried out in the framework of Delphine Traber’s thesis [24], show a 33.68% long-term increase in the PANAS score among members of the Bataillon de Chasseurs Alpins who had undergone training (pretraining score=12.32, posttraining score=15.21, and 6-month posttraining score=16.47). Thus, in the PREPAR study, we consider a 20% improvement as significant (for questionnaire details, see Multimedia Appendix 2 [42-44]).

Self-Compassion
According to Kotsou and Ley [44], the mean score is 2.88 in the French population. A score between 3.5 and 5.0 indicates a high level. Thus, in the context of this study, an increase of 20% (0.62) will be considered relevant (for questionnaire details see Multimedia Appendix 2).

Hardiness, the Ability to Stay Healthy Under Stressful Conditions
In the literature, global hardness scores (measured by the DRS-15 scale) in military populations are around 29 (scores can range from 0 to 45). To the best of our knowledge, there are no studies that have evaluated the effect of a prevention program on this variable. Thus, as this is a dispositional measure, we consider a 5% increase as relevant (for questionnaire details see Multimedia Appendix 2).

Secondary End Points
The evaluation criteria used to meet the secondary objectives are provided in the following sections.

Objective 2.1: Vulnerability
We will assess the following regarding vulnerability:

1. Innate: based on a study of the polymorphism of genes involved in stress regulation mechanisms (see Multimedia Appendix 3). These analyses will not be used in a diagnosis.
2. Acquired: based on miRNAs that target regulatory phenomena established in earlier work (see Multimedia Appendix 4).
Objective 2.2: Change in Resilience
We will assess follow-up of the change in scores on composite end point questionnaires: PANAS, Self-Compassion Scale, and DRS-15 at 6 and 12 months after inclusion.

Objective 2.3: Change in Psychobiological Biomarkers of Activation and Deterioration
We will assess the following regarding change in physiobiological biomarkers of activation and deterioration:
1. Indirect markers of oxidative stress: lipoperoxidation markers thiobarbituric acid reactive substances and 8-iso-prostaglandin F2alpha [45,46].
2. Circulating markers of central nervous system activity: GABA, brain-derived neurotrophic factor, kynurenic acid, and dopamine [47-49].
3. Inflammation: proinflammatory cytokines (including C-reactive protein, TNF-alpha, IL23, and IL12), anti-inflammatory cytokines (IL10 and IL6), and chemokines [50,51].
4. Hypothalamic-pituitary-adrenal axis: cortisol (analysis of saliva on waking, getting up, and 30 minutes later), catecholamines, and neuropeptide Y [52].
5. Markers of physiological activation (autonomic nervous system): cardiac variability index (temporal, frequency, and nonlinear analysis of the electrocardiogram signal to measure indices of parasympathetic flexibility), and electrodermal conductance (level of tonic activity, and amplitude of phasic activity to assess activation of the sympathetic system and its persistence) [53,54].
6. Psychological: follow-up of perceived stress (Cohen questionnaire) and psychological symptomatology scores (the PCL-5, the HADS, and the BMS).

Objective 2.4: Change in Psychosocial Factors
We will assess the following regarding change in psychosocial factors:
1. Flexibility of coping strategies, Flex Cop [55].
2. Sense of duty, inner correspondence, and peaceful harmony (ICPH) questionnaire [31].
3. Emotional reactivity (scores on the emotional Stroop task are compared with the classic Stroop task before and after the intervention at 4 and 12 months) [56,57]. The aim of this task is to evaluate the flexibility of sympathetic and parasympathetic systems inherent in the emotional response to trauma. The classical Stroop task (baseline) will be followed by the emotional Stroop task (reactivity). No learning effect has been documented with using this experimental modality.
4. Self and public stigma scores of PTSD in the military.
5. Perceived social support (the multidimensional scale of perceived social support) [33], and institutional support (the Sieblod vertical and horizontal cohesion questionnaire).

Objective 2.5: Describe Changes in Adherence to the Intervention
Adherence will be assessed using quantitative criteria at 4, 6, and 12 months (the Treatment Motivation Questionnaire). This analysis will be based on (1) an analysis of the processes put in place to operationalize the intervention; (2) a focus group consisting of 8 participants, and 10 individual interviews carried out at 4 and 12 months after inclusion; and (3) a qualitative analysis of discussions that took place during workshops to anchor the practice held at 6 months and 1 year. The latter will also make it possible to identify both obstacles and drivers of adherence.

Objective 2.6: Describe User Satisfaction With the URGOfeel System
We will assess user satisfaction with the URGOFeel (UrgoTech) application using visual analog scales, and short questionnaires that measure (1) the user experience and (2) its benefits.

Ancillary Study
We plan to run a qualitative analysis of the transferability of the intervention. The transferability assessment will be carried out at the very end of the research protocol (1 year after inclusion) and will use the transferability and support to the adaptation of health promotion interventions ASTAIRE tool. The aim is to facilitate the transfer of the intervention to other populations, both military and civilian. The intervention could then be tailored to other populations, based on this evaluation. The ancillary study will focus on qualitative details, notably the conditions, obstacles, and drivers facilitating the transferability of the intervention to other populations or other contexts. The criteria required by ASTAIRE will be supplemented by a qualitative evaluation based on interviews (n=20), and a focus group (n=8) at the end of the intervention. To limit bias, we will ensure that participants in the various segments of this ancillary study (ASTAIRE, focus groups, and interviews) are distinct from one another. A summary of the analysis of these criteria (qualitative, Astra, and quantitative, in the long term) could be used as a guide for future deployments of the intervention.

Statistical Methods
Continuous variables will be presented as mean and SD, if the distribution is normal (the Shapiro-Wilk test will be used, if necessary). In case of a nonnormal distribution, data will be presented as median, quartiles, and extreme values. Scale variables will be treated as ordinal data and analyzed as above. Depending on the analysis, data may be classed as categorical variables. Categorical variables will be expressed as absolute values and percentages. To control for attrition, calculations will assume 20% of data are missing.

The primary analysis will examine the effectiveness of the prevention program. This will be evaluated by comparing the percentage of responses between intervention and nonintervention groups, using a χ² test. Secondary analyses will examine the aforementioned markers. Values for the intervention and the nonintervention group will be compared using a repeated measures ANOVA if the conditions for applying an ANOVA are met. Vulnerability will be initially identified using a cluster analysis of the measured biological variables (a k-means or Gaussian model mixture, depending on the distribution). The analysis will seek to identify the most biologically at-risk cluster (vulnerable compared with nonvulnerable). The effect of the intervention as a function of vulnerability status will be evaluated using repeated measures.
ANOVA, by comparing vulnerable and nonvulnerable groups. Psychosocial variables will be compared between intervention and nonintervention groups using repeated measures ANOVA, if the conditions for its application are met. A regression analysis will aim to test the mediating role of psychosocial determinants in the improvement of the composite score.

Sample Size
The number of participants to be included is not based on an a priori calculation. This is due to the absence of data in the literature regarding the effectiveness of a primary prevention program in the military context, measured using a composite criterion that encompasses affect, self-compassion, and hardiness. In order to determine the sample size needed to observe a significant between-group difference, we used the following parameters: a 2-tailed statistical test; a 55% response rate in the intervention group; a 30% response rate in the nonintervention group (significance $P=.50$; power =0.80). If we take into account potential dropouts (estimated at 20%), each group should consist of 58 individuals, making a total of 116 participants.

Results
The project is currently ongoing, and results are expected to be published by the end of 2024.

Discussion
This project aims to evaluate an interventional research program for the primary prevention of occupational health problems. The program is as comprehensive as possible and synergistically links physiological, psychological, and social factors. We hope that the integration of biopsychosocial factors, which are suited to the characteristics of the occupational environment, will reinforce the effectiveness of current strategies. The project respects the three pillars of evidence-based prevention: (1) it investigates what is the best evidence and contributes to research-based knowledge; (2) it considers experiential knowledge; and (3) it takes into account the values, preferences, and characteristics of populations and individuals. This type of primary prevention intervention could provide a framework for other interventions that can be modified and adapted to different professional contexts. Data regarding transfer and feasibility collected during our study with operational military personnel could be used in further work to optimize the program for other army corps. Finally, it is imperative to stress that this study adopts an ecological approach, that is, it strives to reflect as closely as possible the real-life experiences of French army soldiers. This approach is essential to obtain information directly applicable to their daily routines and challenges. However, it is essential to recognize that there are potential limitations inherent in this approach, particularly with regard to adherence to the research protocol. Operational constraints within the military environment can sometimes override the ideal execution of the research program. These constraints may require adjustments to the protocol to ensure practicality and feasibility. For example, modifications in the number and frequency of workshops may be necessary to meet service needs and requirements, without compromising operationality. However, the encouraging exploratory results and the interest shown by the military personnel observed by Traber [24] mean that we can be confident that this study will provide a better understanding of the field of PTSD prevention in the military.

The research team is committed to maintaining the integrity of the study while adapting flexibly to these challenges, in order to obtain meaningful and relevant results that will contribute to the well-being and effectiveness of our soldiers in the French army.

Acknowledgments
The authors warmly thank officers and operational staff of the Mountain Infantry Brigade and, more specifically, the 4th company of the 3rd Batallon de Chasseurs Alpins and the 3rd company of the 7th Bataillon de Chasseurs Alpins. This work is supported by the French Direction Générale de l’Armement. The opinions or assertions expressed herein are the private views of the authors and are not to be considered as official, or as reflecting the views of the French Military Health Service.

Data Availability
The data sets generated during or analyzed during this study are not publicly available due to French Army data policy but are available from the French Military Health Service at dcssa-paris@sante.defense.gouv.fr upon request.

Authors’ Contributions
SP, ELB, ETC, MH, MT, DC, and AMD were involved in the conception and design of the trial, and they were also responsible for obtaining ethics committee approval. SP, ELB, MMB, MH, MT, DC, and AMD wrote the paper. All the authors contributed to the refinement of the study protocol and approved the final version of the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Workshop schedule.
References


27. Perceived Stigma and Barriers to Care for Psychological Treatment: Implications for Reactions to Stressors in Different Contexts. US: Guilford Publications; 2008.


Abbreviations

- BMS: Burnout Measure Short Version
- DRS-15: Dispositional Resilience Scale
- HADS: Hospital Anxiety and Depression Scale
- ICPH: inner correspondence and peaceful harmony
- MSPSS: Multidimensional Scale of Perceived Social Support
- OPEX: external operation
- PANAS: Positive and Negative Affect Scale
- PCL-5: Posttraumatic Checklist-5
- PHQ-15: Patient Health Questionnaire
- PREPAR: Primary Prevention of Posttraumatic Stress Disorder in Military Professionals
- PSS: Perceived Stress Scale
- PTSD: posttraumatic stress disorder
- TE: traumatic event
Implementation of a Primary Prevention Program for Posttraumatic Stress Disorder in a Cohort of Professional Soldiers (PREPAR): Protocol for a Randomized Controlled Trial


Please cite as:
Implementation of a Primary Prevention Program for Posttraumatic Stress Disorder in a Cohort of Professional Soldiers (PREPAR): Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2024;13:e47175
URL: https://www.researchprotocols.org/2024/1/e47175
doi:10.2196/47175
PMID:38277204

©Emmanuelle Le Barbenchon, Marion Trousselard, Sonia Pellissier, Mathilde Moisseron-Baudé, Philippine Chachignon, Pierre Bouny, Emma Touré Cuq, Sandrine Jacob, Cécile Vigier, Maud Hidalgo, Damien Claverie, Anais M Duffaud. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.01.2024. 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.
Investigation of the Association Between e-Cigarette Smoking and Oral Mucosal Health Status Among Young People: Protocol for a Case-Control Trial

Siyuan Cheng¹, MD
Massachusetts College of Pharmacy and Health Sciences, Boston, MA, United States

Corresponding Author:
Siyuan Cheng, MD
Massachusetts College of Pharmacy and Health Sciences
179 Longwood Avenue
Boston, MA, 02115
United States
Phone: 1 617 732 2850
Email: m0496197@stu.mcphs.edu

Abstract

Background: Given the paucity of current safety studies related to e-cigarettes, there are no definitive studies on whether e-cigarettes cause oral mucosal lesions or even oral cancer. Although it is still undetermined whether e-cigarettes are harmless, an increasing number of teenagers choose to smoke e-cigarettes and believe that they are not harmful to the human body.

Objective: This aims to determine whether e-cigarettes cause damage to the oral mucosa. This study also aims to evaluate the association between e-cigarette smoking and oral mucous membrane lesions in young adults. The objectives are to (1) compare the oral mucosal conditions in participants with and without e-cigarette smoking habits, (2) assess the effect of the amount of e-cigarette smoking on oral mucosal conditions, and (3) assess the effect of the duration of e-cigarette smoking on oral mucosal conditions.

Methods: In this prospective study, 304 youths aged 15 to 24 years (n=152, 50% who smoke only e-cigarettes and n=152, 50% who do not smoke e-cigarettes or cigarettes) will be divided into 2 groups for a controlled study. Whether e-cigarettes cause oral mucosal lesions will be verified by comparing the odds of oral mucosal lesions in the 2 experimental groups. For this experiment, the predefined power is 80% (P=.04), and the predefined proportions of groups 1 and 2 are 11% and 2.5%, respectively.

Results: This experiment is at the conceptualization phase and has not yet been carried out. Experimenters have not been recruited and no data have been collected.

Conclusions: e-Cigarettes are still an unfamiliar topic to the public, and it is still unknown whether they can cause damage to the oral mucosa. This experiment aims to find out whether there is a link between the 2. There are still many limitations in this study, such as the lack of categorization of e-cigarettes and the lack of testing methods for oral mucosal status. These limitations are expected to be addressed in the future as the experiment is formally conducted and further optimized.

International Registered Report Identifier (IRRID): PRR1-10.2196/53644

KEYWORDS
oral mucosal lesions; e-cigarette; youth; oral; moth; lesion; lesions; cigarette; cigarettes; smoker; smoking; smokers; smoke; mucosa; mucosal; dental; dentist; dentistry

Introduction

Background

Oral cavity cancer is the most prevalent head and neck malignancy worldwide [1,2]. This malignant phenotype is often associated with habitual and lifestyle factors, such as tobacco smoking, excessive alcohol consumption, betel nut chewing, and low intake of fruits and vegetables [1]. Due to restrictive government policies and proven negative health effects in many parts of the world, the use of tobacco has declined in recent decades [3]. Electronic cigarettes (e-cigarettes) were invented by a Chinese pharmacist, Hon Lik, in 2003, who envisioned that they would replace conventional cigarettes due to their...
deleterious effects [4]. e-Cigarette companies claim that handheld devices can provide smokers with the same experience as conventional cigarettes while reducing their negative effects. e-Cigarettes were introduced with the hope that the smoking population would gradually stop using conventional cigarettes and switch to e-cigarettes. With the waning consumption of regular cigarettes, the use of e-cigarettes has surged worldwide, indicating that smokers now consider e-cigarettes viable replacements. Moreover, recent data have indicated that e-cigarette smoking practices are more common among teenagers and young adults [5,6]. However, most people do not take this soaring statistic seriously and allow young people to use e-cigarettes freely. Most people believe that the health hazards of smoking only manifest with increasing age and, thus, concern is unnecessary regarding the use of e-cigarettes by young people.

It is no surprise that young adults are the primary habitual e-cigarette users. The portability of the devices, different flavors with less nicotine, and convenient use of e-cigarettes make them appealing to young people. Data from the 2011-2018 National Youth Tobacco Survey in the United States demonstrated that e-cigarette use among high school students increased from 1.5% in 2011 to 20.8% in 2018 [3]. In 2022, the Centers for Disease Control and Prevention indicated that 2.55 million US middle and high school students reported current (past 30 days) e-cigarette use; nearly 85% of these young people used flavored e-cigarettes, and more than half used disposable e-cigarettes [7]. This increase has also become noticeable in community settings. Unfortunately, adults do not pay enough attention to the use of e-cigarettes, especially in schools.

Current research on e-cigarettes has solely focused on their ingredients and compared satisfaction levels of e-cigarettes with those of regular cigarettes [8,9]. The safety of e-cigarettes, especially regarding the etiology of oral and maxillofacial diseases or other possible intraoral side effects, is still unclear. Moreover, few studies have reported on the adverse effects of e-cigarette smoking [10]. Recent experiments have only indicated that the use of e-cigarettes could impact the balance in the oral microbiome while allowing for the rapid growth of foreign microorganisms [4]. Nevertheless, many studies recommend that individuals replace cigarettes with e-cigarettes irrespective of concerns about their safety.

Smoking causes damage to the oral mucosa as well as lesions that can lead to the progressive development of oral cancer [11]. According to research, smoking can cause “oral mucosal leukoplakia” and a variety of other oral mucosal diseases. Additionally, it has been shown that leukoplakia is the oral mucosal lesion that is most likely to lead to oral cancer [12]. Most of the white spots on the lips occur on the lower lip, in the junction of the middle, and on the outer third of the lip, which is where people usually hold cigarettes; this can explain the relationship between holding cigarettes and white spots on the lip. White spots are precancerous lesions, and approximately 4% to 7% can develop into oral cancer. According to statistics, 93.1% of patients with oral white spots are smokers [13]. The development of oral mucosal lesions is based on the principle that tobacco irritates the oral mucosa (through cigarette smoke and the chemicals in cigarettes), causing it to react adversely to prolonged irritation. e-Cigarettes operate in a similar way to tobacco, which also irritates the oral mucosa through atomization and chemicals.

The only difference between e-cigarettes and cigarettes is simply that e-cigarettes do not contain tar; both contain many chemicals as well as nicotine. e-Cigarettes tend to have more types of chemicals than cigarettes because of the wide variety of flavors. Given that it is still unknown whether e-cigarettes influence the development of oral mucosal lesions, the possibility that e-cigarettes can cause damage to the oral mucosa and increase the risk of oral cancer similar to tobacco needs to be studied.

To date, no definite investigation of the association between e-cigarettes and oral cavity cancer has been conducted. When available, this study will be the first to determine whether e-cigarettes are mechanistically linked to the development of oral cancer to further educate young users and students on the potential for malignancy due to e-cigarette use. It is assumed that e-cigarettes cause damage to the oral mucosa and oral cancer, which is similar to cigarettes. The following experimental studies and designs will be conducted to test this hypothesis. Therefore, this study aims to investigate the association between e-cigarette smoking and oral mucosal health status among young adults and determine whether e-cigarettes cause damage to the oral mucosa and increase the risk of oral cancer. Compared to that of other similar studies, the methodology of this study is simpler, and the results will be clearly comparable. If the results show that e-cigarettes can cause damage to the oral mucosa similar to that caused by cigarettes, this study could provide strong support for subsequent experiments to determine whether e-cigarettes cause oral cancer.

**Hypothesis**

This study hypothesizes that e-cigarettes, in the speculated absence of tobacco, may still induce carcinogenesis. Alternatively, it is suggested that this habit may trigger an imbalance in the oral microbiome that predisposes the mucosal epithelium to oral cavity cancer with other etiological factors. This presupposed risk is likely proportional to the frequency and duration of e-cigarette smoking. The oral mucosal health of participants will be determined by determining the presence of white spots in the mouth (the main feature of the oral mucosa that is damaged by smoking behavior).

**Primary Outcome**

The primary outcome is the effect of e-cigarette smoking on oral mucosal conditions.

**Research Significance**

The results of this study will provide preliminary evidence of the malignant potential of e-cigarettes in the deterioration of oral mucosal conditions or exclude them as a cause of oral carcinogenesis. If a direct association is observed, the results of this study would be vital to inform e-cigarette users of the harmful and even carcinogenic nature of e-cigarette smoking, despite e-cigarettes being considered tobacco free. In addition, this would help guide the implementation of legislative policies to bring awareness to the additional risks associated with the use of e-cigarettes and vaporizers. Moreover, these findings, if
positive, will pave the way for future research on chemical products that may be present in e-cigarettes that are directly involved in the malignant transformation of oral keratinocytes.

Methods

Trial Registration
Given that this study is at this stage of conceptualization and has received no support or sponsorship from any organization at this time, this experiment has not yet been applied for trial registration. The registration will be done in the future when support is received from the relevant organizations.

Ethical Considerations
This study has not yet been submitted to the ethical review board for assessment, mainly due to lack of financial support and lack of assistance from large organizations. We emphasize that this is an independent study and is still at the research protocol stage. As the project has not received sufficient financial support, we are unable to cover the costs of applying for evaluation by the ethics review committee at this time. At this stage, we are working to ensure that the study design meets ethical standards and will seek possible review and approval at a future stage. We understand and value participant rights and will take appropriate measures to ensure the ethical and legal nature of the study.

Study Design
The adoption of a case-control, prospective, observational study design is motivated by the relatively low incidence of oral mucosal lesions within the general population. This design allows for a focused exploration of the relationship between e-cigarette smoking and oral mucosal health.

Case-control design is particularly suitable for investigating rare outcomes such as oral mucosal lesions, as it efficiently compares individuals with the outcome of interest (cases) to those without (controls).

A prospective design involves following participants over time, allowing for the collection of data on exposures and outcomes as they occur. By prospectively tracking participants, the study can gather real-time information on e-cigarette smoking habits and observe the development of oral mucosal lesions, providing a temporal sequence crucial for establishing causation.

An observational design is chosen over an experimental one due to ethical considerations and the nature of the research question. Since randomly assigning participants to smoke e-cigarettes for an extended period is ethically challenging, an observational approach allows for the examination of naturally occurring exposure to e-cigarette smoking in real-world settings.

The low incidence of oral mucosal lesions in the general population necessitates a design that efficiently targets and investigates this specific outcome. By focusing on a population at risk (e-cigarette smokers) and carefully selecting controls, the study maximizes its ability to detect and understand the potential impact of e-cigarette smoking on oral mucosal health.

Study Population
According to the United Nation definition, people between the ages of 15 and 24 years are defined as young people [14]. This age group coincides with the rapidly growing population of e-cigarette smokers, which is not a concern to mainstream society. Therefore, the 15- to 24-year-old group has been selected as the study population. American participants between 15 and 24 years of age will be recruited irrespective of their sex, race, occupation type, and socioeconomic status and divided into “e-cigarette case” and “control” groups. The selection criteria for each group are as follows.

The exclusion criteria for the two groups (the following criteria will be applicable in both groups and do not need to be listed repeatedly) are (1) participants with histologically diagnosed recurrent oral cavity cancer whose primary tumors occurred outside the recruitment timeframe (2 years before the first diagnosis); (2) participants with synchronous solid or hematological malignancies in other regions at the time of oral cavity cancer diagnosis; (3) participants with a genetic predisposition to oral cavity cancer, including those with Fanconi anemia, systemic lupus erythematosus, and dyskeratosis congenita; (4) participants who meet the inclusion criteria but are unwilling to participate in the study after detailed information has been provided; and (5) participants with severely debilitating systemic conditions that preclude participation in the study.

Selection of the “e-Cigarette Case” Group

Inclusion Criteria
A prospective study among experimental participants who smoked e-cigarettes over the last 210 days or 7 months will be conducted. According to the article “Effects of Duration of Electronic Cigarette Use,” the average duration of use among e-cigarette smokers is 210 days or 6.8 months [15]. The same criteria will be used in this study to standardize the duration of use for both the conventional cigarette smoking population and the e-cigarette smoking population.

Exclusion Criterion
The exclusion criterion is participants who smoke tobacco.

Selection of Controls
Inclusion criteria are individuals without a history of smoking e-cigarettes, drinking alcohol, and smoking tobacco at the time of recruitment.

Exclusion criteria are participants who smoke e-cigarettes and tobacco.

Recruitment Strategy

Overview
The recruitment strategy, targeting customers at e-cigarette stores and pedestrians, is strategically aligned with the study’s focus on e-cigarette users. The inclusion of community selection for matching controls, if quotas are not met, demonstrates flexibility in the recruitment approach. The disclosure of convenience and nonprobability sampling methods ensures transparency, thereby acknowledging the limitations inherent
in these methods. A nuanced understanding of participant selection challenges is crucial for the accurate interpretation of study findings.

Community Selection Details
If quotas for controls are not met through the initial recruitment strategy, a community selection approach will be considered. This involves identifying and recruiting controls from community settings, ensuring a diverse representation.

Clinical Recruitment Procedures

e-Cigarette Recruitment Procedures
The e-cigarette store recruitment consists of 3 stages, which are as follows:

1. Approach: e-cigarette stores will be approached to seek their cooperation in the recruitment process.
2. Informed consent: store owners and managers will be provided with detailed information about the study objectives, procedures, and ethical considerations. Upon agreement, consent forms may be obtained from the store owners to allow recruitment on their premises.
3. Participant identification: e-cigarette users within the specified age range will be approached, and the study will be explained to them. Interested individuals will be given detailed information about the study and their consent will be sought.

Pedestrian Recruitment
The pedestrian recruitment consists of 2 stages, which are as follows:

1. Approach: pedestrians in high-traffic areas will be approached with an invitation to participate in the study.
2. Informed consent: similar to the e-cigarette store recruitment, detailed information will be provided to potential participants, and consent will be obtained before proceeding with any study-related activities.

Community Selection (if Necessary)
If quotas for controls are not met through the initial recruitment strategy, community selection may be considered.

1. Approach: community settings such as local community centers or public spaces will be identified. Consent from relevant authorities and community leaders will be sought.
2. Participant identification: controls meeting the study criteria will be approached in these settings and recruitment will follow the same informed consent procedures.

Sampling Methods
Convenience and nonprobability sampling methods will be used for the recruitment of cases and controls. The detailed sample size determination is shown in Multimedia Appendix 1 [16]. Briefly, the predefined power is 80% (P<.05), and the predefined proportions of the case and controlled groups are 11% and 2.5%, respectively. In all, 304 youths aged 15 to 24 years (n=152, 50% who smoke only e-cigarettes and n=152, 50% who do not smoke e-cigarettes or cigarettes) will be recruited.

Data Collection: Questionnaire
Interviewer-administered questionnaires will be used to collect data on demographic, lifestyle, and socioeconomic variables for the cases and controls. The questionnaire tool will comprise 2 parts: exposure information (part A) and sociodemographic information (part B). A life grid sheet will be used during interviews to efficiently collect retrospective information on lifestyle habits and other risk factors. To reduce observer bias, interviewers will not initially be informed about the aim of the research. Once cases are identified and interviewed, matching controls will be recruited and interviewed. Answers provided by participants will be converted to objective scores after data collection by the investigator (SC). The details on the parts of the questionnaire are as follows.

Exposure information (part A) will include (1) oral mucosal health status (whether white spots are present), (2) smoking habits, (3) type of products used, (4) e-cigarette smoking status (yes or no and current or previous), (5) age of onset, (6) duration of smoking, (7) pattern of smoking duration (continuous or intermittent), (8) frequency of smoking (daily, weekly, and occasionally), (9) number of cartridges (frequency), and (10) Tobacco consumption. Sociodemographic information (part B) will include (1) age, (2) sex, (3) occupation type, (4) education level, (5) income, and (6) ethnicity.

Questionnaire Validity and Reliability
Before the deployment of the questionnaire for field use, face validity and qualitative content validity will be determined by a panel of experts in oral and maxillofacial surgery, dental public health, biostatistics, and laypersons. The internal consistency of the scales used to measure e-cigarette smoking exposure will be determined using Cronbach α, with a minimum α value of .70 indicating good and acceptable reliability.

Data Analysis
Descriptive statistics will be used for all binary, categorical, and continuous variables and expressed as tables, texts, and figures. Bivariate analysis will then be performed for relevant variables. The Shapiro-Wilk test will be performed for continuous variables to determine whether they follow the Gaussian distribution. Afterward, the independent 1-tailed t test and 1-way ANOVA will be used; otherwise, the Mann-Whitney U test and Kruskal-Wallis test will be performed. For cross-tabulation of 2 categorical variables, the chi-square test will be used. Variables that do not fulfill the assumption of this test will be analyzed using Fisher exact test. Comparisons with P<.20 will be used to implement multivariable analysis using multiple logistic regression. Odds ratios and 95% CIs will be determined for e-cigarette smoking and other individual factors. For all analyses, P<.05 will be used to denote statistical significance. Statistical analyses will be performed using the SPSS (version 27; IBM Corp).

The control and e-cigarette groups will be subjected to a t test (assuming the data meet the normality criteria and do not meet the criteria for the rank sum test) to verify whether e-cigarettes cause harm.
Results

This experiment is at the conceptualization phase and has not yet been carried out. Experimenters have not been recruited and no data have been collected.

Discussion

In examining the potential negative impact of e-cigarettes on the oral mucosa, this study deliberately focused its investigation on this specific aspect, excluding other potential confounding factors that might influence the results. Notably, variables such as alcohol consumption, daily routines, and gender differences were not considered in the experimental design. While these factors were omitted from this study, they were duly documented in the questionnaire phase, laying the groundwork for future investigations to delve into these variables. This strategic choice aimed to isolate the primary relationship between e-cigarette use and oral mucosal health, with the intention to refine and expand the scope in subsequent research endeavors.

The findings of this study align with existing literature on e-cigarettes, which suggests a correlation between the chemicals present in e-cigarettes and cytotoxicity, leading to damage to the oral mucosa. This corroborates with prior research, providing further evidence of the potential harm associated with e-cigarette use [10,16,17]. The recognition of this association holds significant implications for public health, emphasizing the need for awareness campaigns and regulatory measures to mitigate the adverse effects on oral health. Moreover, the study acknowledges the diverse landscape of e-cigarettes in the United States, encompassing various brands and flavors, each containing distinct chemical compositions. Given this complexity, it is acknowledged that the study does not comprehensively analyze all individual chemicals present in different e-cigarettes. This limitation prompts a call for future experiments to undertake a more nuanced examination of the diverse chemical profiles within e-cigarettes to better understand their distinct impacts on oral mucosal tissues.

Despite the valuable insights provided by this study, certain limitations should be acknowledged. The decision to exclude factors such as alcohol consumption and daily routines may have implications for the generalizability of the findings. In addition, the complexity of e-cigarette compositions poses a challenge, as the study did not extensively investigate the myriad chemicals present in different e-cigarette products. Future research endeavors should consider a more comprehensive approach, encompassing a broader range of variables and a detailed analysis of the chemical constituents of various e-cigarettes.

In conclusion, this study contributes valuable evidence to the growing body of knowledge on the potential negative impact of e-cigarettes on oral mucosal health. By focusing on a specific aspect while recognizing its limitations, this research paves the way for future investigations to build upon these findings. The implications extend beyond the immediate scope of oral health, emphasizing the broader need for public health interventions and regulatory measures in response to the evolving landscape of e-cigarette use.

Data Availability

Data sharing is not applicable to this paper as no data sets were generated or analyzed during this study.

Authors' Contributions

SC was in charge of the study conceptualization, writing and reviewing the original draft, and editing the final draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample size determination.

Data Availability

Data sharing is not applicable to this paper as no data sets were generated or analyzed during this study.

Authors' Contributions

SC was in charge of the study conceptualization, writing and reviewing the original draft, and editing the final draft.

Conflicts of Interest

None declared.

References

6. Teens are 16 times more likely to use JUUL than older age groups. Truth Initiative. 2018. URL: https://truthinitiative.org/research-resources/emerging-tobacco-products/teens-are-16-times-more-likely-use-juul-older-age [accessed 2024-01-05]


Leveraging Community Health Workers and a Responsive Digital Health System to Improve Vaccination Coverage and Timeliness in Resource-Limited Settings: Protocol for a Cluster Randomized Type 1 Effectiveness-Implementation Hybrid Study

Lavanya Vasudevan1,2*, CPH, MPH, PhD; Jan Ostermann2,3*, PhD; Nathan Thielman2,4, MD, MPH; Joy Noel Baumgartner5, PhD; David Solomon6, MD; Anna Mosses6, MD; Amy Hobbie3, MPH; Nicole L Hair3, PhD; Chen Liang3, PhD; Marco van Zwetselaar7, MSc; Sayoki Mfinanga6, MD, PhD; Esther Ngadaya6*, MD, PhD

1Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States
2Duke Global Health Institute, Duke University, Durham, NC, United States
3Department of Health Services Policy & Management, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States
4Department of Medicine, Duke University, Durham, NC, United States
5School of Social Work, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
6National Institute for Medical Research, Muhimbili Research Centre, Dar es Salaam, United Republic of Tanzania
7Zwets IT, Harskamp, Netherlands
*these authors contributed equally

Corresponding Author:
Jan Ostermann, PhD
Department of Health Services Policy & Management
Arnold School of Public Health
University of South Carolina
915 Greene St, Discovery 351
Columbia, SC, 29208
United States
Phone: 1 8037778747
Email: jano@mailbox.sc.edu

Abstract

Background: Tanzania is 1 of 20 countries where the majority of unvaccinated and undervaccinated children reside. Prior research identified substantial rural-urban disparities in the coverage and timeliness of childhood vaccinations in Tanzania, with children in rural settings being more likely to receive delayed or no vaccinations. Further research is necessary to identify effective and scalable interventions that can bridge rural-urban gaps in childhood vaccination while accounting for multifaceted barriers to vaccination.

Objective: This protocol describes a type 1 effectiveness-implementation hybrid study to evaluate Chanjo Kwa Wakati (timely vaccination in Kiswahili), a community-based digital health intervention to improve vaccination timeliness. The intervention combines human resources (community health workers), low-cost digital strategies (electronic communication, digital case management, and task automation), a vaccination knowledge intervention, and insights from behavioral economics (reminders and incentives) to promote timely childhood vaccinations.

Methods: The study will be conducted in 2 predominantly rural regions in Tanzania with large numbers of unvaccinated or undervaccinated children: Shinyanga and Mwanza. Forty rural health facilities and their catchment areas (clusters) will be randomized to an early or delayed onset study arm. From each cluster, 3 cohorts of mother-child dyads (1 retrospective cohort and 2 prospective cohorts) will be enrolled in the study. The timeliness and coverage of all vaccinations recommended during the first year of life will be observed for 1200 children (n=600, 50% intervention group children and n=600, 50% nonintervention group children). The primary effectiveness outcome will be the timeliness of the third dose of the pentavalent vaccine (Penta3). Quantitative surveys, vaccination records, study logs, fidelity checklists, and qualitative interviews with mothers and key informants will inform the 5 constructs of the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework. The results will be used to develop an implementation blueprint to guide future adaptations and scale-up of Chanjo Kwa Wakati.

https://www.researchprotocols.org/2024/1/e52523
Results: The study was funded in August 2022. Data collection is expected to last from February 2024 to July 2027.

Conclusions: This study will address the lack of rigorous evidence on the effectiveness of community-based digital health interventions for promoting vaccination coverage and timeliness among children from sub-Saharan Africa and identify potential implementation strategies to facilitate the deployment of vaccination promotion interventions in low- and middle-income countries.

Trial Registration: ClinicalTrials.gov NCT06024317; https://www.clinicaltrials.gov/study/NCT06024317

International Registered Report Identifier (IRRID): PRR1-10.2196/52523

(JMIR Res Protoc 2024;13:e52523) doi:10.2196/52523

KEYWORDS

childhood vaccinations; timeliness; vaccine hesitancy; digital health; community health workers; Tanzania; low- and middle-income countries; SMS; reminder; conditional incentive

Introduction

Background

Globally, the number of children missing their first dose of the diphtheria-pertussis-tetanus (DPT) vaccine rose from 19 million in 2019 to 25 million in 2021 [1,2]. The vast majority (14.2 million, 78%) of such zero-dose children reside in 20 countries, including Tanzania [3]. As in other low- and middle-income countries (LMICs), substantial rural-urban disparities in routine childhood vaccination exist in Tanzania, with vaccination rates being lower in rural areas than in urban areas [4]. Using data from the 2015-2016 Tanzanian Demographic and Health Survey, we demonstrated gaps in vaccination coverage (receipt of each recommended vaccine dose by age 1 y) and timeliness (receipt of each vaccine dose within 28 d of the recommended age) for Tanzanian children [4]. The coverage of the first dose of the pentavalent vaccine (Penta1, which includes antigens of DPT, Hemophilus influenzae, and hepatitis B) was 79.4% nationally in this analysis, with the remaining 21% of children classified as zero dose [4]. We documented receipt of the third dose of the pentavalent vaccine (Penta3) in even fewer children (72.7%), suggesting dropouts in the multidose vaccine series. Finally, rural children had lower timeliness of vaccination (47.8% delayed for Penta3) than urban children (24.2% delayed for Penta3). New interventions are needed to reduce the number of children who are zero dose, receive delayed vaccines, or drop out and to close the rural-urban gap in vaccination [4,5]. Such interventions must consider multifaceted barriers to vaccination and variations in the availability of human resources and infrastructure in rural areas.

In prior research in southern Tanzania, we identified challenges with the availability of, and access to, vaccination services, including challenges with distance and transportation to health facilities, temporary nonavailability of vaccines owing to a lack of reliable refrigeration at health facilities, and vaccine wastage policies that prevented the use of multidose vials when clinics had low volumes of children to be vaccinated [6]. Service unreliability and lack of communication about service interruptions were noted as causes of frustration among mothers of vaccine-eligible children [6]. Our research also identified vaccine hesitancy owing to vaccine-related knowledge gaps and concerns. In general, rural mothers reported more vaccine-related knowledge gaps and concerns than urban mothers [6]. Despite challenges, vaccination intention was high among mothers, and conditional incentives were identified as a potential nudge to increase vaccination timeliness [7,8]. In addition, a community health worker (CHW)–delivered knowledge intervention was piloted and determined to be feasible to implement in rural areas. On the basis of the findings of this formative research and other published reports [9], we designed Chanjo Kwa Wakati (timely vaccination in Kiswahili, the most commonly spoken language in Tanzania), a community-based digital health intervention to improve childhood vaccination coverage and timeliness. The intervention seeks to combine human resources (CHWs), low-cost digital strategies (electronic communication, digital case management, and process automation), a vaccination knowledge intervention (counseling scripts addressing specific knowledge gaps), and insights from behavioral economics (reminders and incentives) to promote timely childhood vaccination.

Objectives

This protocol describes our planned evaluation of the implementation and impact of Chanjo Kwa Wakati. The study seeks to contribute evidence to the literature in 3 key ways. First, although many studies have evaluated mobile phone–based reminders for promoting childhood vaccinations [10], evidence is lacking [9] for more complex community-based digital health interventions that target multifaceted barriers to vaccinations, such as those identified in our prior research [4,6,7]. Second, few community-based digital health interventions for promoting childhood vaccination have been evaluated using rigorous randomized controlled study designs, especially in sub-Saharan African countries and in rural areas [11-15]. Third, evidence on implementation strategies associated with deploying community-based digital health interventions in LMICs is limited but critically important for supporting scale-up in the context of highly resource-constrained national health systems. To bridge these gaps in the literature, we will use a type 1 effectiveness-implementation hybrid study [16,17] to evaluate the effectiveness of Chanjo Kwa Wakati in improving the timeliness of childhood vaccinations in rural areas in Tanzania and identify strategies that support its implementation. The detailed protocol of the study is presented in the following sections.
Methods

Study Overview and Aims
This type 1 effectiveness-implementation hybrid study [16,17] uses a cluster randomized trial to evaluate intervention effectiveness and mixed methods to describe implementation outcomes. A cluster randomized design was chosen because the delivery of the intervention is organized at the cluster level. In addition, there are ethical concerns with the randomization of intervention activities, including incentives, at the individual level in rural communities in Tanzania. The evaluation of the Chanjo Kwa Wakati intervention will be guided by the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework. The methods used in this study are described in accordance with the CONSORT (Consolidated Standards of Reporting Trials) checklist for cluster randomized studies (Multimedia Appendix 1). The study aims and hypotheses are described in the following subsections.

Aim 1
The first aim is to evaluate the effectiveness of Chanjo Kwa Wakati for increasing the timeliness of childhood vaccinations due by age 1 year compared with the standard of care. The effectiveness of Chanjo Kwa Wakati will be evaluated in a cluster randomized trial with 1200 mother-child dyads enrolled from the catchment areas of 40 rural health facilities (clusters) in 2 regions in Tanzania. Clusters will be randomized into an early or delayed onset study arm. The intervention will target the mother-child dyad; outcomes will be assessed at the child level. Our hypothesis is that Chanjo Kwa Wakati is effective for increasing the timeliness of childhood vaccinations due by age 1 year compared with the standard of care. The primary outcome measure used to evaluate intervention effectiveness will be a continuous measure of the timeliness of the third dose of the pentavalent vaccine (Penta3), due at age 14 weeks, and determined using the birth and vaccination dates abstracted from official vaccination cards issued for the child. A secondary outcome will be a binary measure of timeliness, defined as receipt of Penta3 within 28 days of the due date. Other outcomes measures include the timeliness and coverage of all other vaccine doses due by age 1 year.

Aim 2
The second aim is to evaluate the implementation factors associated with variation in intervention effectiveness and develop an implementation blueprint for intervention scale-up to other settings. Study logs, fidelity checklists, quantitative surveys, and qualitative interviews with mothers and key informants will be used to inform other constructs of the RE-AIM framework, specifically reach, adoption, implementation, and maintenance. Two key implementation outcomes of interest are participants’ reports of the receipt of intervention components (intervention fidelity) and their reports of intervention acceptability. Analyses of variation in intervention implementation at the cluster level and systematic variation in the effectiveness for aim 1 outcomes across children will guide the future optimization of Chanjo Kwa Wakati. The results will be used to develop an implementation blueprint to guide future adaptations and scale-up of Chanjo Kwa Wakati.

Study Setting
The study will be conducted in 4 rural districts in Mwanza and Shinyanga regions in Tanzania. Both regions have large numbers of unvaccinated or undervaccinated children. The populations of Mwanza and Shinyanga region were estimated to be 3,699,872 and 2,241,299, respectively [18]. In 2022, only 56% of children aged 12 to 23 months in Mwanza region and 32.2% of children in Shinyanga region were estimated to have received all basic vaccinations, including one dose of Bacillus Calmette-Guerin vaccine, 3 doses of polio vaccine, 3 doses of DPT-containing vaccine, and one dose of measles-containing vaccine [19]. Fewer children aged 12-23 months were considered fully vaccinated according to the national vaccination schedule shown in Table 1 (32.8% in Mwanza and 2.5% in Shinyanga) [19]. Notably, 3.3% of children aged 12-23 months in Mwanza and 14.5% in Shinyanga were reported to have received no vaccinations. [19].

Table 1. Routine childhood immunizations recommended in Tanzania before age 1 year.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Age at vax due date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillus Calmette-Guerin (BCG) and oral polio vaccine (OPV) 0</td>
<td>At birth or first contact</td>
</tr>
<tr>
<td>OPV1; pentavalent vaccine comprising antigens of diphtheria, pertussis, tetanus, <em>Haemophilus influenza</em> B, and hepatitis B (Penta) 1; pneumococcal vaccine (PCV) 1; and rotavirus vaccine (Rota) 1</td>
<td>6 weeks</td>
</tr>
<tr>
<td>OPV2, Penta2, PCV2, and Rota2</td>
<td>10 weeks</td>
</tr>
<tr>
<td>OPV3, Penta3, PCV3, and injectable polio vaccine (IPV)</td>
<td>14 weeks</td>
</tr>
<tr>
<td>Measles and rubella 1</td>
<td>9 months</td>
</tr>
</tbody>
</table>

Intervention Components
Chanjo Kwa Wakati focuses on individual-level barriers to timely vaccination to complement the Tanzanian government’s efforts to strengthen vaccination systems and reduce structural barriers [20]. Chanjo Kwa Wakati comprises the following individual-level intervention components: a vaccination knowledge intervention, reminders about upcoming vaccination due dates, service notifications, and conditional incentives for timely vaccination (Table 2). The intervention’s target population is the mother-child dyad from the time of pregnancy (third trimester) until the child is aged 1 year. Intervention activities are targeted toward the mother, whereas the timeliness of Penta3 (primary outcome) and other vaccinations recommended before age 1 year (secondary outcomes; Table 1) is assessed for the child.
Table 2. The Chanjo Kwa Wakati intervention.

<table>
<thead>
<tr>
<th>Timing</th>
<th>Activities</th>
<th>Detailed description of activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last trimester of pregnancy</td>
<td>• Enrollment&lt;br&gt;• Registration in digital health system&lt;br&gt;• Knowledge intervention&lt;br&gt;• Baseline assessment&lt;sup&gt;a&lt;/sup&gt;</td>
<td>• Informed consent&lt;sup&gt;a&lt;/sup&gt;&lt;br&gt;• Register pregnancy and phone numbers in the digital health system&lt;br&gt;• Identify knowledge gaps for an individualized knowledge intervention based on a study by Vasudevan et al&lt;sup&gt;[6]&lt;/sup&gt;&lt;br&gt;• Assess postintervention change in vaccination knowledge and attitudes as well as potential correlates of coverage and timeliness&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>≤4 wk after the date of delivery</td>
<td>• CHW&lt;sup&gt;b&lt;/sup&gt; phone call or home visit</td>
<td>• Ascertain date and place of birth as well as the receipt of vaccinations due at birth&lt;br&gt;• Register the birth, record the 6-wk vaccination due date, and update phone numbers in the digital health system</td>
</tr>
<tr>
<td>Before each vaccination due date (6, 10, and 14 wk, and 9 mo)</td>
<td>• SMS text message reminders with incentive offers&lt;br&gt;• Backup: CHW phone calls or home visits&lt;br&gt;• Knowledge intervention</td>
<td>• Messages are sent to the mother and vaccine advocates 7 d and 1 d before each vaccination&lt;br&gt;• Messages are individualized and account for information on service interruptions&lt;br&gt;• Messages include conditional incentive offers&lt;br&gt;• Individualized knowledge intervention focused on vaccinations due at an upcoming visit</td>
</tr>
<tr>
<td>≤1 wk after each vaccination due date</td>
<td>• CHW phone call or home visit</td>
<td>• Verify vaccination status and date against vaccination card and issue incentive as appropriate&lt;br&gt;• Update vaccination due dates (and thus reminders) as necessary</td>
</tr>
<tr>
<td>12-15 mo after birth</td>
<td>• Follow-up assessment&lt;sup&gt;a&lt;/sup&gt;</td>
<td>• Validate vaccination coverage and dates using vaccination cards&lt;sup&gt;a&lt;/sup&gt;&lt;br&gt;• Issue incentive as appropriate</td>
</tr>
</tbody>
</table>

<sup>a</sup>Research activity. All other activities are intervention activities.

<sup>b</sup>CHW: community health worker.

### Individualized Knowledge Intervention

The knowledge intervention consists of 9 true or false statements that assess key knowledge gaps about childhood vaccinations. A brief counseling script is provided with each statement that will be read out by the CHW after an incorrect response. Statements are based on prior research in Tanzania [6]. Statements and counseling scripts were tested extensively for content, comprehensibility, and cultural appropriateness with national key informants from Tanzania’s Immunization and Vaccines Development program, regional and district health officials, CHWs, and pregnant women from rural Tanzania. In addition to the in-person individualized vaccination education provided by CHWs, SMS text message–based vaccination promotion messages, sent via the Chanjo Kwa Wakati digital health system, will be individualized to each participant’s knowledge gaps and aligned with the child’s vaccination schedule.

### Vaccination Reminder Messages

CHWs will verify the due date for the next vaccination using the child’s vaccination card and update the child’s information in the Chanjo Kwa Wakati digital health system. The system will automatically schedule SMS text message reminders to be sent 7 days and 1 day before each vaccination due date. The SMS text messages will include the name of the intended recipient, the child’s name, the vaccination due date, and other relevant information.

### Service Notifications

In the event of stockouts or service nonavailability, SMS text messages with relevant information will be sent to mothers whose children are due for vaccination.

### Incentive Offers

Conditional incentive offers will be specific to each child’s vaccination schedule. In our prior work [7,8], we identified a set of incentives (eg, pharmacy vouchers and birth certificates) that are likely to be acceptable to mothers. The specific incentives, their value, conditionality, and the timing of disbursement will be determined in discussion with local key informants.

### Implementation Strategies

Relevant implementation strategies will be identified in collaboration with local key informants, and a detailed logic model will be developed to guide study activities.

### Intervention Activities

Figure 1 describes the sequence of intervention activities. Pregnant women, enrolled in their third trimester, will receive a home visit from a CHW before the birth of their child, be registered with the Chanjo Kwa Wakati digital health system, and receive an individualized knowledge intervention to screen and counsel for common knowledge gaps and concerns about routine childhood vaccinations. After the birth of the child, the
mother will receive individualized SMS text message reminders for upcoming vaccinations, messages aimed at mitigating persistent vaccination knowledge gaps, service notifications (eg, related to stockouts or service nonavailability), and conditional economic incentive offers aimed at encouraging timely vaccinations. CHWs will use the digital health system for work planning, recording vaccination dates and due dates, and following up on missed vaccination appointments and backstop the digital health system with phone calls and in-person visits as needed for vaccination reminders and assessing and addressing mothers’ knowledge gaps.

Figure 1. Sequence of intervention activities. CHW: community health worker.

Study Population

Sampling Area
The sampling area comprises 2 rural districts in Mwanza region and 2 rural districts in Shinyanga region, selected based on district-level information about vaccination rates, logistical considerations, and input from key informants.

Cluster Eligibility
Twenty clusters will be selected in each region for a total of 40 clusters. Clusters are defined as the catchment area of eligible health facilities in the sampling area. Eligible health facilities include public and public-designated (private not-for-profit facilities that serve the functions of public health facilities) hospitals, health centers, and dispensaries. Facilities must be operational, offer routine prenatal care and childhood vaccination services, have at least 2 active CHWs operating in the catchment area, and have reported at least 100 pregnancies or births in the year before study implementation. Eligibility will be determined initially using administrative data obtained from district offices and verified using surveys with health facilities.
Study Participants, Eligibility, and Recruitment

The study population will include 3 groups of participants.

Cross-Sectional Retrospective Cohort

This cohort (n=400) will be used for retrospective assessments of children’s vaccination records during their first year of life. Eligible participants will be mothers or legal guardians (henceforth referred to as *mothers*) of children aged 12 to 23 months, aged ≥15 years, and residing in the sampling area since the birth of the child.

Before recruiting this cohort, CHWs, with help from local key informants (eg, village leaders, health care providers, and traditional birth attendants), will compile a list of children aged 12 to 23 months residing in each cluster. The lists will be randomized, and mothers will be approached by the research team for eligibility determination, informed consent, and enrollment until 10 mother-child dyads from each cluster are enrolled into the study. Up to 40 additional women, who are not necessarily living in the study area but meet the other eligibility criteria, may be enrolled to pilot-test the study instruments.

Longitudinal Prospective Cohort

This cohort (n=800) will be used for prospective assessments of children’s vaccination records during their first year of life. Eligible participants will be pregnant women in their last trimester of pregnancy, aged ≥15 years, residing in the sampling area, and expected to reside in the sampling area until the child reaches age 1 year.

Before recruiting this cohort, CHWs, with help from local key informants (eg, village leaders, health care providers, and traditional birth attendants) will compile a list of pregnant women residing in each cluster. The lists will be randomized, and pregnant women will be approached by the research team for eligibility determination, informed consent, and enrollment until 10 pregnant women from each cluster are enrolled into the study. One additional eligible woman will be enrolled from each cluster during each prospective enrollment round to account for loss to follow-up (refer to the *Retention* subsection).

Participants in Qualitative Feedback

This group will be used to obtain feedback on the RE-AIM constructs. Eligible participants will be key informants at the national, regional, or local levels (target: n=12), health providers from participating facilities who are responsible for childhood vaccinations (target: n=40, approximately 1/cluster), CHWs in study clusters (target: n=80, approximately 2/cluster), and women enrolled in the prospective cohort (target: n=60).

To recruit these participants in qualitative work, a combination of purposive and snowball sampling strategies will be used. The numbers of participants represent target numbers; additional participants will be enrolled if saturation is not reached. Key informants, including policy makers, decision makers, and implementers at the national, regional, or local levels (eg, officials from Tanzania’s Immunization and Vaccines Development program as well as regional, district, and local health officials) will be asked to participate in qualitative interviews on implementation factors and to recommend other individuals who can provide relevant information. Health providers who are responsible for childhood vaccinations will be enrolled from participating health facilities. CHWs will be enrolled from participating clusters. Women will be purposively selected from the prospective cohort based on data collected during enrollment or follow-up surveys, study logs, or the digital health system during study implementation.

Study Design

Cluster Randomized Trial Design

Forty rural health facilities and their catchment areas (*clusters*), 20 per region, will be randomized to an (A) early or (B) delayed onset study arm (*Figure 2*). From each cluster, 3 cohorts of 10 mother-child dyads (1 retrospective cohort and 2 prospective cohorts) will be enrolled into the study. For the retrospective cohort, 400 mothers of children aged 12 to 23 months (10/cluster) will be enrolled as a cross-sectional sample for retrospective assessments of children’s vaccination records during the first year of life. These cohorts, denoted by A1 and B1, will be part of the care-as-usual control group. For the longitudinal prospective cohort, 800 pregnant women in their last trimester of pregnancy will be enrolled during 2 rounds of enrollment (10 women/enrollment round/cluster) for prospective assessments of their children’s vaccination records during the first year of life. These cohorts are denoted by A2, A3, B2, and B3. In the first round of enrollment, 200 (50%) of the 400 prospective participants (those enrolled from early onset clusters; A2) will receive the intervention, and the other 200 (50%) participants (B2) will be part of the care-as-usual control group.

In the second round of enrollment, all 400 participants (A3 and B3) will receive the intervention. In total, vaccination uptake and timeliness will be observed for 1200 children (n=600, 50% intervention group children and n=600, 50% nonintervention group children) during their first year of life. The nonintervention group children represent the care-as-usual control group in the trial.

The trial design, which analytically resembles a difference-in-difference design, allows for an analysis of the effectiveness of the intervention even in the context of contemporaneous changes in the rate and timeliness of vaccinations at the national level (eg, through national vaccination campaigns or changes in the vaccination schedule).
Figure 2. Design of the 2-arm cluster randomized trial (n=1200).

**Evaluation Framework**

The evaluation of the Chanjo Kwa Wakati intervention will be guided by the RE-AIM framework. Quantitative survey data will be the primary source of information for the analysis of the effectiveness of the intervention. These data will be complemented by study logs, fidelity checklists, and qualitative interviews with Tanzanian key informants, facility-based health workers, CHWs, and mothers to assess the acceptability, reach, and fidelity of Chanjo Kwa Wakati, as well as factors that may inform sustainability and scalability in the future.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type</th>
<th>Outcome measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaccination cards</td>
<td>Primary</td>
<td>Timeliness of the third dose of the pentavalent vaccine, Penta3 (number of days delayed)</td>
</tr>
<tr>
<td>Vaccination cards</td>
<td>Secondary</td>
<td>Timeliness of Penta3 (≤28 days delayed)</td>
</tr>
<tr>
<td>Vaccination cards</td>
<td>Other</td>
<td>Timeliness of all other vaccine doses recommended by age 1 y (number of days delayed)</td>
</tr>
<tr>
<td>Vaccination cards</td>
<td>Other</td>
<td>Coverage of all vaccine doses recommended by age 1 y</td>
</tr>
</tbody>
</table>

**Study Procedures**

**Cluster Randomization**

In each district, all eligible health facilities will be assigned random IDs. Random IDs will be generated using a random number generator in Stata (version ≥16; StataCorp LLC) with a fixed seed to ensure reproducibility. In each district, the 10 clusters with the smallest random IDs will be allocated to either the early onset arm (odd ranks) or delayed onset arm (even ranks). Facilities that are not willing to participate or are deemed ineligible will be replaced with the next odd- or even-ranked facility, respectively. If the distributions of key characteristics of clusters (eg, the number of pregnant women/y, the number of births/y, the percentage of home births, vaccination coverage, and the number of days vaccinations are offered each wk) differ significantly between study arms, a covariate-constrained randomization approach will be used to maximize the balance of these characteristics across study arms.

**Outcomes Assessments**

Vaccination outcomes are described in Table 3. Vaccination coverage and dates will be ascertained for each study child using government-issued vaccination cards, which will be photographed, scanned, or copied and entered by research staff. For the retrospective cohort, the enrollment survey will assess vaccination outcomes for the child; their participation will end after the enrollment survey. Outcome data for the prospective cohort will be collected during an endline survey 12 to 15 months after the birth of the child. All vaccination doses of all vaccination series will be tracked. For women who cannot provide vaccination cards at the time of the survey, consent will be sought to access their children’s paper-based or electronic vaccination records at their local health facility.

**Blinding**

Owing to the nature of the intervention, there will be no blinding of investigators, implementers (CHWs), data collectors, or participants.

**Enrollment**

Participants may be enrolled at their homes, health facilities, or other mutually agreed-upon locations. The potential risks and benefits of research participation will be carefully explained during the informed consent process (refer to the Ethical Considerations subsection). To maximize the reach of the intervention, participants without mobile phones will be able to designate a vaccine advocate of their choice to receive mobile phone–based messages.

**Retention**

Retention only applies to the prospective cohort. To ensure the retention of study participants, extensive contact information, including the names and mobile phone numbers of both parents, vaccine advocates, and other contacts, as well as expectations for the place of delivery, travel plans after birth, and GPS
coordinates of homes will be documented at enrollment. In the event of child death, maternal death, or participant withdrawal, intervention activities related to the mother-child dyad will be suspended, and the dyad will be withdrawn from the study. In the event of child relocation or family travel during study activities, efforts will be made to collect outcomes data for the child using a combination of mobile phone–based surveys; a review of the child’s records at the local health facility; and proxy reports from the father, other legal guardians, or other informants. As the intervention involves local CHWs, for participants relocating outside their cluster for the remainder of the study period, intervention activities will stop after relocation.

Data Collection

Overview

Trained research assistants will conduct enrollment and data collection in Kiswahili. Survey instruments will be developed in English, translated into Kiswahili, and back-translated into English. The development of the surveys will be informed by prior surveys [6,8].

Enrollment Survey

All participants will complete an enrollment survey. The survey will last approximately 45 minutes and assess knowledge and attitudes, sociodemographic characteristics of participants and their households, access barriers, digital literacy, residual knowledge gaps, and other correlates of vaccine uptake and timeliness. The mother’s engagement in prenatal care will be assessed via self-report and scanned prenatal care cards. For retrospective participants, the enrollment survey will also assess vaccination uptake and timeliness (Table 3), whereas for prospective participants, extensive contact information will be collected (refer to the Retention subsection).

Follow-Up Survey

Prospective participants will complete a follow-up survey 12 to 15 months after their child’s birth. The survey will last approximately 30 minutes. The primary purpose of the follow-up survey is to collect data on vaccination outcomes (Table 3). The follow-up survey will include the knowledge intervention that addresses each woman’s vaccination-specific knowledge gaps and assess vaccination attitudes and experiences with vaccine uptake and timeliness during the study period. The follow-up survey will also be used to update information on mothers’ engagement in prenatal care via self-report and scanned prenatal care cards. For intervention group women, the follow-up survey will also assess implementation measures, experiences with the intervention, and the acceptability and perceived efficacy of Chanjo Kwa Wakati.

Facility Survey

During the enrollment and follow-up work, we will conduct a survey of the health facility in each cluster. The purpose of the facility survey is to assess the implementation context. The survey will collect service information (eg, information on vaccine stockouts and facility closures), procedural knowledge (eg, vaccination dates, scheduling, and rescheduling), and contact information.

Qualitative Data

In-depth interviews (IDIs) will elucidate wide-ranging feedback on all constructs of the RE-AIM framework. IDIs will be conducted by local trained study staff in English or Kiswahili. Study staff will follow a semistructured interview guide. IDIs will be recorded to facilitate transcription and data analysis. Depending on the availability of the participants, IDIs may be conducted by phone or in person. IDIs are expected to last 45 to 60 minutes.

Statistical Power

The aim 1 cluster randomized trial will include 1200 women, enrolled from 40 different clusters, across 3 distinct enrollment phases, with 10 women enrolled from each cluster in each phase. This equates to k=120 statistical clusters with n=10 women per cluster. Half of these clusters (k=60/120, 50%; n=600/1200, 50%) will be allocated to the intervention arm and half to the control arm. Assuming that the primary outcome—the timeliness of the Penta3 vaccine—has an SD of 28 days, the power of the trial to detect an intervention-related difference of 7 days is 0.95 with an intracluster correlation coefficient of 0.05; the power is 0.88 with an intracluster correlation coefficient of 0.10.

The aim 2 qualitative formative work will yield data from IDIs with key informants (n=12), facility-based health workers (up to n=40), CHWs (up to n=80), and mothers (up to n=60). Prior research suggests that thematic saturation can be achieved with 12 interviews, with metathemes presenting as early as 6 interviews [23-25]. Thus, these proposed sample sizes should be sufficient to achieve thematic saturation of implementation barriers and residual barriers to timely and equitable vaccinations, as well as to identify any regional variation in reasons for delayed vaccinations.

Data Analysis

Quantitative Analysis of Intervention Effectiveness

The primary outcome measure of interest will be a continuous measure of vaccination timeliness, expressed as the delay, in days, between the vaccination due date and the date on which the vaccination was received for the third dose of the pentavalent vaccine, Penta3, due at age 14 weeks (Table 1). The secondary and other outcome measures include the timeliness and coverage of all vaccine doses recommended by age 1 year. Owing to the randomization, a comparison of mean outcomes between participants in the intervention group and those in the nonintervention group yields an unbiased estimate of the effect of the intervention.

To control for variation in sociodemographic characteristics across participants, seasonal effects, and other potential correlates of vaccination timeliness, outcomes will also be analyzed in a multivariable regression framework using accelerated failure time models with the completed vaccination dose as the failure event, age in days describing time to failure, and membership in the intervention versus control groups as the primary covariate of interest.

Robust SEs will be estimated to account for the clustering of observations across mother-child dyads enrolled from the same geographic clusters at the same time.
Analysis of Implementation Outcomes
We will generate summary measures (eg, means) for continuous data and use proportions to summarize categorical data. For outcomes assessed using validated measures (eg, the Acceptability of Intervention Measure [26]), we will present composite scale scores. Qualitative implementation outcomes will be analyzed as described below.

Analysis of Qualitative Data and Integration With Quantitative Data
Thematic analyses will be facilitated by qualitative software (eg, NVivo [Lumivero]) and a codebook made up of a priori and emergent structural codes based on the interview guide and 4 interrelated steps: reading, coding, data display, and data reduction [27]. The data will be used to fully map the Chanjo Kwa Wakati process from the perspectives of the mother and the CHW and characterize the opportunities and limitations of using digital and in-person resources to support vaccination timeliness. Using convergent mixed methods, we will use joint displays and narrative integration to connect the quantitative and qualitative data [28-30].

Analysis of Systematic Variation in Effectiveness
Extensive sensitivity analysis for the aim 1 primary and secondary outcomes analyses will characterize variation in intervention effectiveness with population, setting, and implementation factors. Variation in intervention effectiveness will be evaluated using interactions between geographic, maternal, or child characteristics and the intervention variable. Statistically significant parameters on interaction terms are indicative of differential intervention effects for different subgroups. To account for variation in implementation, sensitivity analysis will include per-protocol (as-treated) and intention-to-treat analyses in which a vector of variables describing the fidelity of the different intervention components will be used in place of the binary intervention variable.

Missing Data
The Study Limitations and Adaptations subsection in the Discussion section discusses considerations regarding missing data on outcomes and covariates and their handling in the analyses.

Ethical Considerations
The study protocol, informed consent forms, and procedures have been reviewed and approved by the Health Sciences South Carolina Institutional Review Board (the reviewing institutional review board [IRB]: Pro00120675) and the ethics review committee of the National Institute for Medical Research in Tanzania (NIMR/HQ/R.8a/Vol.IX/4242), as well as the relying IRBs at Duke University (Pro00117772), Emory University (STUDY00005518), and the University of North Carolina at Chapel Hill. Informed consent will be sought from all study participants.

To maintain confidentiality, interviewers will conduct the consent process in a private space or in a mutually agreed-upon location where other people are not present. The interviewers will read out the consent form to potential participants and answer any questions. The potential risks and benefits of participation as well as the details of data confidentiality and use will be carefully explained in culturally appropriate and understandable language. Participants can refuse to participate and will only be consented after the research staff is satisfied that they understand all substantive provisions of the informed consent document. Persons with obvious psychological or psychiatric disorders that preclude informed consent will be excluded. Participants will be asked to provide written consent. They will be provided with a copy of the consent form. Informed consent documents will be developed in English, translated into Kiswahili, and back-translated into English.

Participants in the trial will be aged ≥15 years and either pregnant in their last trimester of pregnancy or mothers or legal guardians of children aged 12 to 23 months. For adolescents aged 15 to 17 years who are pregnant or mothers of children aged 12 to 23 months and are otherwise eligible to participate in the study, we obtained a waiver of parental consent. Tanzanian national policy indicates that adolescents aged <18 years who are sexually active and pregnant have the right to access reproductive health services (eg, HIV testing, care, and treatment [31], as well as family planning [32]) without parental or spousal consent, and they can make associated medical decisions on their own behalf. Given that these adolescents have this adult right to medical decision-making, we posit that they are also able to consent to research participation for themselves without parental or spousal consent. Risks from participation in this study are considered commensurate with ordinary daily life.

Protocol amendments will be submitted to the relevant IRBs for approval before implementation. Adverse events will be reported to the ethics review committee of the National Institute for Medical Research in Tanzania and the US-based IRBs.

Participants will be compensated for the time they spend participating in study activities. Ethically and culturally appropriate compensation will be established in collaboration with community advisers and other key informants. Subject to IRB approval and discussions with key informants, incentive amounts may be adjusted during the study period to account for exchange rate fluctuations and cost-of-living increases.

Results
Results are pending. The study was funded in August 2022. Data collection is expected to last from February 2024 to July 2027.

Discussion
Relevance
The Chanjo Kwa Wakati intervention and proposed evaluation are aligned with the strategic priorities of the Immunization Agenda 2030 and Tanzanian priorities for digital health and community health workforce development [33,34]. Specifically, our study builds on a Tanzanian digital health investment road map [35] to digitize health care data and is in line with Tanzania’s 2020 National Operational Guideline for Community-Based Health Care Services, which seeks to build a strong community health workforce for health promotion,
including for child health [36]. The relevance of our study extends beyond rural Tanzania. The low-cost strategies developed and evaluated in this study are applicable to other rural contexts where some populations are less likely to be vaccinated, experience poor vaccine access, or express low confidence in vaccinations, and our study may inform the development and implementation of multipronged programs to promote timely vaccinations in such settings.

**Contributions to the Literature**

Our study is expected to make several important contributions to the scientific literature on digital health and vaccination interventions. First, numerous studies have documented low timeliness of vaccinations for children from LMICs, with a disproportionate burden for rural children [4,6,37-41]. Our study is significant because it seeks to evaluate an intervention for promoting vaccination timeliness for rural children. Second, lay health care workers such as CHWs play a critical role in delivering health services in LMICs. Our study will use a hybrid digital and in-person approach and contribute evidence on effective and scalable strategies for supporting lay health workers with digital health tools to promote childhood vaccination equity. Third, although many studies have evaluated mobile phone–based reminders for promoting childhood vaccinations [10], evidence is lacking for community-based systems that target multifaceted barriers to vaccinations. Our study will contribute data on the effectiveness of a community-based digital health intervention to promote equitable and timely vaccination in LMICs. Fourth, research studies support the use of screening tools to identify vaccination knowledge gaps and deliver targeted interventions to combat vaccine hesitancy. Our study will integrate CHW outreach and mobile phone–based vaccination education to bridge gaps in vaccination knowledge, respond to parental concerns, and reduce hesitancy. Finally, our study will contribute evidence on strategies to bridge the digital divide in rural areas and the gender gap in mobile phone access by using a combined digital and in-person approach.

**Study Limitations and Adaptations**

**Unknown Number of Eligible Women per Cluster**

Although we estimate that 20 clusters per region are required to reach our enrollment targets for the cross-sectional and longitudinal samples, neither the number of eligible women per cluster nor the number of refusals are known a priori. If the sample size across clusters is too small, sampling may be extended to additional clusters until sample size targets have been reached in each region and study arm.

**Randomization**

If the distributions of key characteristics of clusters vary greatly between the initially specified study arms (based on randomized IDs), a covariate-constrained randomization approach may be used that maximizes the balance across study arms in the distribution of key characteristics of clusters.

**Spillover and Contamination**

Cluster randomization is expected to minimize spillover effects; however, contamination may exist for participants living at the boundaries between intervention and nonintervention clusters. Sensitivity analyses will evaluate potential spillover effects by including geographic proximity to an intervention cluster as a covariate for the nonintervention group women in our models.

**External Events**

A potential threat to validity stems from the possibility of a local vaccination campaign or selective stockouts among subsets of clusters during follow-up. Given our collaboration with regional and national stakeholders, enrollment from multiple clusters, multiple vaccination time points for each participant, and the robust design, the likelihood and potential impact of such external events on our findings is low. The effects of stockouts on timeliness will be assessed in the per-protocol sensitivity analysis.

**Missing Data**

If vaccination cards are missing for >10% of the cross-sectional sample, and their vaccination information cannot readily be ascertained from records at the local health facility, enrollment into the cross-sectional cohort will continue until key outcome measures can be ascertained for the target number of 400 mothers. This may include the expansion of enrollment to additional clusters. Enrollment into the longitudinal cohort will be increased proportionately. Missing data on other characteristics will be addressed using full information maximum likelihood estimation or sensitivity analyses [42-45].

**Mobile Phone Access and Use**

In the target population, mobile phone access and use is not universal, and mobile phones may be shared within households. Chanjo Kwa Wakati is specifically designed to compensate for gaps in mobile phone ownership and coverage, including through the designation of vaccine advocates who will receive reminders and in-person outreach by CHWs to those who are not reached by mobile phone.

**Sustainability of Reminders and Incentives**

Reminder messages and incentive offers, which were informed by our prior work, will be carefully reviewed by investigators and key informants to ensure their ethical implementation and sustainability for a scaled-up intervention.

**Generalizability**

To assess the generalizability of the study’s context and findings, study data and the relationships among key variables of interest will be compared with information in publicly available survey data and administrative data sources.

**Interpretation and Dissemination**

If successful, this study will contribute data on the effectiveness and implementation of the Chanjo Kwa Wakati intervention that engages recent mothers to increase the timeliness and coverage of routine childhood vaccinations in rural Tanzania. Key findings will be presented in peer-reviewed manuscripts, presentations at national and international conferences, and media outlets. Policy and programmatic recommendations will be developed in collaboration with key informants and decision makers at the national, regional, and district levels.
Acknowledgments

The authors would like to thank the Duke Global Health Institute Research Design & Analysis Core staff for their input on cluster randomization and sample size calculations. The research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (R01HD110844). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The data from this study will be available from the National Institute for Medical Research, Muhimbili Research Centre, Dar es Salaam, Tanzania, but restrictions apply that are governed by consent forms and data transfer agreements. Data may be made available by the authors upon reasonable request and with the permission of the National Medical Research Review Committee of the National Institute for Medical Research, Tanzania. The open-source digital health system will be made available to interested researchers and practitioners.

Authors' Contributions

LV, JO, and EN designed the study with critical input from JNB, NT, DS, MvZ, and AH. LV and JO wrote the first draft, and EN, JNB, NT, DS, MvZ, and AH contributed to revisions. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT (Consolidated Standards of Reporting Trials) 2010 checklist of information to include when reporting a cluster randomized trial.
[DOCX File, 36 KB - resprot_v13i1e52523_app1.docx ]

Multimedia Appendix 2
Peer-review report: Clinical Informatics and Digital Health Study Section, Healthcare Delivery and Methodologies Integrated Review Group, Center for Scientific Review, National Institutes of Health.
[PDF File (Adobe PDF File), 155 KB - resprot_v13i1e52523_app2.pdf ]

References


Abbreviations

CHW: community health worker
CONSORT: Consolidated Standards of Reporting Trials
DPT: diphtheria-pertussis-tetanus
IDI: in-depth interview
IRB: institutional review board
LMICs: low- and middle-income countries
OPV: oral polio vaccine
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance
Testing an Evidence-Based Self-Help Program for Infertility-Related Distress: Protocol for a Randomized Controlled Trial

Jennifer L Gordon¹, PhD; Megan M L Poulter¹, BA; Ashley A Balsom¹, MA; Tavis S Campbell², PhD

¹Department of Psychology, University of Regina, Regina, SK, Canada
²Department of Psychology, University of Calgary, Calgary, SK, Canada

Corresponding Author:
Jennifer L. Gordon, PhD
Department of Psychology
University of Regina
3737 Wascana Parkway
Regina, SK, S4S 0A2
Canada
Phone: 1 306 585 4389
Email: jennifer.gordon@uregina.ca

Abstract

Background: Infertility—the inability to achieve pregnancy despite ≥12 months of focused attempts to conceive—is experienced by 1 in 6 couples. Women typically carry a disproportionate share of the psychological burden associated with infertility, experiencing poor quality of life, and 30%-40% experiencing depressive mood or anxiety. Unfortunately, currently available psychological interventions targeting infertility-related distress are associated with modest effects.

Objective: Our team, in collaboration with patient advisors, has designed a self-help intervention for infertility-related distress involving 7 weekly 10-minute videos addressing the cognitive, behavioral, and interpersonal challenges associated with infertility, delivered through a mobile app. A feasibility study suggests that it is well accepted and highly effective in reducing symptoms of anxiety and depressed mood among distressed individuals dealing with infertility. This study represents the next step in this line of research: a fully powered randomized controlled trial comparing the intervention to a waitlist control group.

Methods: We will recruit 170 individuals struggling to become pregnant in Canada or the United States to be randomized to our 7-week self-help program or a treatment-as-usual condition. The primary outcome will be fertility quality of life, while secondary outcomes will include depressive symptoms, anxious symptoms, and relationship quality, assessed before and after the program as well as biweekly for 16 weeks following completion of the program. Self-reported health care use and the presence of diagnosed mood and anxiety disorders, assessed through a structured psychiatric interview, will also be assessed immediately following the intervention and at the 16-week follow-up assessment. Treatment adherence and retention will also be recorded throughout the intervention. Multilevel modeling will compare the intervention arm to the treatment-as-usual condition in terms of all continuous outcomes across the 9 measurement points. Logistic regression will be used to assess the occurrence of mood and anxiety disorders in the 2 treatment arms at the posttreatment assessment as well as at the 16-week follow-up. Sensitivity analyses will examine potential treatment moderators: membership in the LGBTQIA+ (lesbian, gay, bisexual, transgender, queer, intersex, and asexual) communities, baseline fertility quality of life, cultural background, disability status, and pursuit of conception through medical intervention.

Results: We expect our intervention to be more effective than treatment-as-usual in improving all mental health parameters assessed and decreasing health care use related to both mental and reproductive health. Effects are expected to be larger with decreasing baseline quality of life and equally effective regardless of membership in the LGBTQIA+ communities, cultural background, disability status, and pursuit of conception through medical intervention.

Conclusions: If our intervention is successful, this would suggest that it should be scaled up and made publicly available. The availability of this program would fill an important gap in light of the high rates of psychopathology among those experiencing infertility and considering the current lack of effective psychotherapy approaches for infertility.

Trial Registration: Clinicaltrials.gov NCT06006936; https://classic.clinicaltrials.gov/ct2/show/NCT06006936
International Registered Report Identifier (IRRID): PRR1-10.2196/52662
Introduction

Overview

The Canadian Community Health Survey reveals that 16% of Canadian reproductive-aged couples are currently infertile [1], defined as being unable to achieve pregnancy despite ≥12 months of focused attempts to conceive. Although male- and female-factor infertility are equally prevalent, women bear the brunt of the infertility-related burden as treatments require that women monitor their menstrual cycles, attend near-daily ultrasounds, self-inject gonadotropins, and undergo invasive and painful procedures. Women who travel for fertility treatments face additional psychosocial burdens, including schedule disruptions, time off work, and coordination of care among multiple health care providers. It is therefore not surprising that women carry a disproportionate share of the psychological burden associated with infertility, with infertile women consistently reporting lower self-esteem, more instances of depressed mood and anxiety, and lower life satisfaction than their male partners [2]. Lesbian couples pursuing sperm donation experience similar distress, with the intended pregnant individual being at higher risk for depression and anxiety relative to their partner [3].

Around 30%-40% of women presenting for the evaluation of infertility report clinical symptoms of depressed mood and/or anxiety [4-7]. In addition, research from our team [8] suggests that the COVID-19 pandemic has exacerbated distress amid fertility treatment suspensions and delays. Untreated symptoms of depressed mood and anxiety among women with infertility may, in turn, reduce the likelihood of achieving pregnancy, given that psychological burden is the most commonly cited reason for prematurely discontinuing fertility treatments [9]. In a study of 450 couples who were offered 3 government-funded in vitro fertilization (IVF) cycles, 54% did not complete all 3 cycles despite not achieving pregnancy, with “psychological burden” being the primary reason for discontinuing IVF [10]. It is critical that women with infertility who are distressed receive effective mental health treatment to reduce distress and improve conception rates.

Despite high rates of distress among women with infertility, currently available psychological interventions are often ineffective or associated with relatively small benefits. In our recent systematic review of psychological interventions for infertility-related distress [11], we observed that typical interventions are associated with a small beneficial effect on anxiety but a nonsignificant effect on depressive mood, marital quality, or quality of life. Our conclusions confirm findings from a previous meta-analysis [12] and those of a recent review [13] concluding that “a new intervention (targeting infertility-related distress) should be developed.”

Systematic Review and Meta-Analysis

As a starting point, we conducted a systematic review and meta-analysis of psychological interventions for infertility [11,16], which included an examination of treatment moderators such as psychotherapeutic approach (eg, mindfulness-based approaches vs cognitive behavioral therapy) and therapy format (eg, self-administered vs group). This process not only confirmed that currently available interventions were largely ineffective but also revealed that neither therapeutic approach nor format significantly impact treatment benefits.

Qualitative Needs Assessment

Our team then used semi-structured interviews with women with infertility and mental health professionals specializing in infertility to identify the unique aspects of infertility-related distress [17]. Table 1 depicts the themes and subthemes identified. Unique features include the avoidance of infertility reminders (eg, pregnant women and children), excessive cognitive and behavioral focus on one’s infertility to the exclusion of previously enjoyed activities, and negative interactions with loved ones perceived as insensitive.
In addition to identifying clear psychological targets for our intervention, this study also aimed to clarify which interventions were currently being used by practicing clinicians. Our findings indicated a near-universal use of an eclectic and unstructured approach associated with the widely held opinion that no existing therapeutic approach sufficiently addresses all of the biopsychosocial factors contributing to infertility-related distress.

**Evaluation of Potential Intervention Components**

The next step in our intervention development was to identify and consider all candidate techniques that might effectively target the psychological challenges identified as being common in infertility. To do so, we identified all of the psychotherapeutic approaches endorsed by the American Psychological Association’s Clinical Section as being evidence-based for the treatment of anxiety, mood disorders, relationship difficulties, and chronic illness. The 5 identified approaches were then further broken down into their component procedures, resulting in a total of 14 different psychological techniques. In collaboration with a panel of patient advisors and using lay language, we described how each of these techniques would look when applied to infertility-related distress and what their purpose would be. We then surveyed a total of 644 women from online infertility-specific support groups [18], asking them to rate the perceived usefulness of each of the 14 techniques while asking them to identify up to 5 that were “most liked” and “most hated.” Participants were also given the opportunity to provide written feedback on each of the techniques, such as how they might be better tailored to infertility. We then presented the results of this survey to our panel of patient advisors and collaboratively decided on the content of our intervention. We decided on 6 core modules plus a bonus module, the content of which is described in Textbox 1 [18].

**Textbox 1.** Chosen modules based on a survey of 644 women with infertility [18] and collaboration with patient advisors. Mean helpfulness ratings for each module, as assessed in a feasibility study of 21 women, are shown.

<table>
<thead>
<tr>
<th>Modules and focus</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive restructuring: identifying and challenging the extreme negative thoughts that contribute to a depressive and anxious mood (eg, “In vitro fertilisation will never work”).</td>
<td>Challenging core beliefs: identifying and challenging unhelpful deep-seated beliefs about themselves, other people, and the world that are perhaps not based on reality (eg, “nothing ever works out for me”). It involves looking for patterns in one’s thinking from the first module.</td>
<td>Behavioral activation: identifying activities that have been dropped or engaged in less because of an increased focus on infertility. Aim to reintegrate these previously enjoyed activities into their day-to-day lives.</td>
<td>Sharing your grief: learning about different styles of coping and how clashes in coping styles can lead to conflict within a couple. The individual is instructed on how to engage their partner in a structured conversation about how each can help the other in times of grief, such as following a negative pregnancy test.</td>
<td>Strengthening your relationship (bonus module): provides evidence-based information about how to better connect with one’s partner in general. Was offered along with Module 4 for those experiencing relationship distress.</td>
</tr>
<tr>
<td>Living your values (ie, stopping avoidance): reflecting on one’s overarching life values and considering how one’s daily actions align with those values. Indirectly addresses avoidance that is common among individuals with infertility (eg, withdrawing from friends and family and avoiding children and pregnant women). Encourages the individual to consider ways in which they can reduce avoidance without worsening their distress.</td>
<td></td>
<td></td>
<td></td>
<td>Living your values (ie, stopping avoidance): reflecting on one’s overarching life values and considering how one’s daily actions align with those values. Indirectly addresses avoidance that is common among individuals with infertility (eg, withdrawing from friends and family and avoiding children and pregnant women). Encourages the individual to consider ways in which they can reduce avoidance without worsening their distress.</td>
</tr>
<tr>
<td>Summary or wrap up: providing an overview of the program’s content and encouraging the individual to reflect on what’s been accomplished as well as areas for further development.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When it came time to decide on the format of our intervention, a self-help internet-delivered intervention was chosen for several reasons. First, the findings from our systematic review indicate that self-administered interventions are as effective in improving mental health as other, more costly formats. Second, when asked what format a new intervention should take, 71% (15/21) of women from our needs assessment qualitative study responded that it should be self-delivered through the internet. Third, we reasoned that this format could be most effectively scaled up and made accessible to diverse populations of women and in...
regions with limited access to psychological services. Thus, in close collaboration with our patient advisors, each module was translated into a 10-minute PowerPoint (Microsoft Corp) slideshow with professional voiceover. A mobile app was then developed to increase accessibility to the modules.

**Preliminary Testing of Our New Infertility-Specific Intervention**

With a new program developed and fully vetted in consultation with our patient advisory panel, we conducted a pilot study of 21 women recruited through local support groups, assessing intervention acceptability [19]. All participants exhibited clinically significant levels of infertility-related distress, as indicated by a Fertility Quality of Life Scale (FertiQoL) score ≤52 [20]. Enrolled through a Zoom (Zoom Video Communications)-administered enrollment session, participants received 1 module per week through email and were asked to view the 10-minute slideshow within 24 hours of receipt. Midweek, they received an email reminder of the homework assignment, encouraging them to apply their homework assignment throughout the week. At the end of each week, participants were asked to rate the extent to which the module and homework were perceived as helpful in lowering their distress (0-10). At baseline and each week, participants completed the Generalized Anxiety Disorder-7 (GAD-7) and 9-item Patient Health Questionnaire (PHQ-9); the FertiQoL was completed at baseline and after the intervention. Each week and in an interview at the end of the study, participants provided written and verbal feedback, respectively, on the intervention.

Of the 21 women enrolled in the study, 2 became pregnant and were removed from the program prematurely (all outcomes assessed until the point of pregnancy were analyzed). Of the remaining 19 women, 15 completed all 6 modules, and 4 completed a portion of the program. Data from all 19 women were included in the analysis. The average helpfulness rating of each module was found to be 7.4 or above out of 10. Fertility quality of life increased by an average of 12 points out of 100, with a Cohen $d$=0.9. Large reductions in both mean symptoms of anxiety and depressed mood were observed (pre-to-post Cohen d=–1.2 and –1.3, respectively, where effects above 0.8 are considered large), corresponding with clinically meaningful improvements. In addition, 85% of participants experienced a clinically significant decline in either anxious or depressive symptoms (defined as a change of 4 points on the GAD-7 [21] and 5 points on the PHQ-9 [22]).

While the intervention was successful, areas for improvement were identified. For example, homework assignments were modified to include examples of completed homework. Participants reported that our “bonus” relationship module deserved to be a core part of the program, extending it to a total of 7 weeks.

**The Current Trial**

Over the last 3 years, our team has carefully designed a self-help intervention for infertility-related distress that is patient-informed and developed using best practices in intervention design. Results from this feasibility study suggest that it is well accepted and effective in increasing quality of life and reducing symptoms of depressed mood and anxiety among women with infertility-related distress. The proposed project, a sufficiently powered randomized controlled trial comparing the intervention to a treatment-as-usual control group, is the next step in this line of inquiry. It is hypothesized that the intervention will result in greater increases in fertility quality of life and relationship quality as well as decreases in symptoms of depressed mood and anxiety relative to a treatment-as-usual condition, and that these improvements will be maintained over a 16-week follow-up assessment period.

If the proposed trial confirms that the intervention is effective in improving quality of life and mental health symptoms among those with infertility, our next step will be to make this program widely available to women, including making the intervention available through YouTube and engaging our collaborating knowledge users and partner organizations to promote it widely. We will also aim to tailor the program for diverse and marginalized underserved groups.

**Methods**

**Trial Design**

The proposed research is a single-blind randomized controlled trial comparing the above-described self-help program to a waitlist control condition. As the project requires no in-person contact, we will recruit women living throughout Canada and the United States. Fertility quality of life, infertility-related distress, symptoms of depressed mood and anxiety, and relationship quality will be assessed before and after the program, as well as every other week for 16 weeks.

**Treatment Conditions**

**Intervention Condition**

Participants will be given access to a 10-minute module video per week through a mobile app created for this trial. Midweek, participants will receive an automated email reminder of the homework assignment for that week, encouraging them to incorporate the homework into their daily lives. Participants will be permitted to engage in any other psychological interventions they wish but will be asked to report other psychological interventions accessed at the end of their participation.

**Treatment-As-Usual Control Condition**

Participants assigned to the control condition will be instructed to continue their pursuits to conceive without accessing the self-help program. They will be permitted to access other psychological resources that are available to them, though, like the intervention condition, they will be asked to report any treatments accessed in the postintervention survey. They will complete the outcome measures at the same time as participants in the treatment condition. Following completion of the study, the control group will be offered the program in the same manner as the treatment group.

In the original funded grant protocol, we had proposed a waitlist control condition in which participants were not permitted to access other mental health services; however, we have since changed the control condition to treatment-as-usual in order to
more accurately estimate the real-world effectiveness of the treatment. This change also allows us to open the trial to individuals reporting suicidal ideation because these participants will likely require additional mental health services while participating in the current trial.

Randomization Scheme
The Clinical Research Support Unit at the University of Saskatchewan will create the randomization scheme and provide the principal investigator with opaque envelopes containing treatment assignments, ensuring that the research team has no control over the assignments. Randomization will take place at the end of each enrollment session, after the baseline surveys have been completed, and will be stratified based on whether a woman is undergoing fertility treatments or attempting to conceive without medical intervention, as this will be a potential moderating variable.

Protecting Against Sources of Bias
A number of strategies will be used to protect against bias. First, the trial will be registered with clinicaltrials.org before any data collection commences. Second, as described above, the randomization scheme will be created by a third party, and the study research assistants will be instructed to strictly adhere to the randomization protocol without exception. Third, though it is not possible to maintain full blinding of either the participant or research team given the nature of the intervention, all outcomes will be collected by a research assistant who is blind to the participant’s treatment allocation. Fourth, an intent-to-treat approach will be taken in analyzing the results—every effort will be made to continue to collect outcome data on all participants, regardless of whether participants dropped out of the intervention early or not. Finally, we will follow the CONSORT-SPI (Consolidated Standards of Reporting Trials for Social and Psychological Intervention Trials) reporting guidelines [23] in reporting the results of the trial, strictly adhering to the original trial protocol. Any deviations will be clearly described and justified.

Participants

Inclusion and Exclusion Criteria
Based on our sample size calculations, we will recruit 170 women, recruited through the web. The inclusion criteria will include the following: is infertile as defined as either actively attempting to conceive for ≥12 months without success or is currently undergoing fertility treatments (e.g., ovulation induction medication, IVF, and intrauterine insemination). This definition ensures that this study is inclusive of both individuals who cannot afford fertility treatments and women who are in same-sex couples and cannot conceive naturally. Though the original funded protocol excluded individuals reporting active suicidal ideation, those already receiving psychotherapy, and those with high levels of fertility-related quality of life (FertiQoL above 70), we have since decided to remove these exclusion criteria in order to closely estimate the program’s anticipated real-world effectiveness. Rather than exclude individuals based on baseline quality of life, we will perform secondary analyses, considering baseline quality of life as a treatment moderator.

Sex and Gender Considerations
In light of research finding that the intended pregnant individual experiences the most distress in the context of infertility, we will only recruit individuals who have a uterus. However, we will ensure that this study is welcoming to individuals of all gender identities and sexual orientations, as this study will aim to contribute to current knowledge surrounding the psychological experiences of individuals from minority genders, and sexual groups experiencing infertility. This study materials including advertisements, will use inclusive language. Advertisements will not use the word “woman” but instead “individuals attempting to get pregnant but experiencing infertility.” The intervention itself has also been designed with inclusive language, a gender-neutral design, and pictures of individuals from diverse backgrounds and sexual orientations. To ensure adequate diversity among our participants, we will advertise on subreddits specifically targeting members of the LGBTQIA+ (lesbian, gay, bisexual, transgender, queer, intersex, and asexual) community.

Participant Screening and Enrollment
Prospective participants will be emailed the link to a web-based eligibility survey. If found to be eligible, they will be asked to provide their contact information, and a research assistant will contact them to schedule an enrollment session through videoconference.

During the Zoom-facilitated remote enrollment session, eligibility will be confirmed, a brief introductory video explaining the study and intervention will be presented, and consent will be obtained. An enrollment session in which visual contact is made will ensure that our recruited participants are not simply “bots” posing as eligible participants. During the session, participants will complete the baseline questionnaires through a link emailed to them by the research assistant. Upon completion of the questionnaires, the research assistant will open an opaque envelope, revealing the participant’s random assignment to either the treatment or control condition. The research assistant will then ask the participant which day of the week they would like to receive their weekly module video (if assigned to the treatment condition) or weekly outcomes survey (if assigned to the control condition).

Participant Safety
Participants endorsing suicidal ideation on question 9 of the PHQ-9 at baseline will be permitted to participate in the study but will be informed that their level of risk will be reassessed weekly. Specifically, each week, they will receive a survey question, “Please pick out the one statement that best described how you have been feeling during the past week, including today: (A) I don’t have any thoughts of killing myself. (B) I have thoughts of killing myself, but I would not carry them out. (C) I would like to kill myself, or (D) I would kill myself if I had the chance.” If participants choose either C or D, a message including contact information for 2 suicide hotlines will appear. As well, an automatic notification will be sent to the study therapist, flagging the response. They will then follow up with the participant immediately by phone, at most within 24 hours.
If participants endorse A or B, they will simply be allowed to continue with the program.

In addition, the presence of active suicidal ideation (presence of a plan or intent) will be assessed by the researcher during the Zoom-facilitated enrollment session. Those endorsing active suicidal ideation will be referred to additional in-person mental health resources available in their geographic area. They will be given access to the program for their own benefit but will not be enrolled in the study.

**Primary and Secondary Outcomes**

**Overview**
Self-reported psychological outcomes will be assessed immediately postintervention (ie, at the end of the 7th week) and every 2 weeks for 16 weeks. Mood and anxiety disorders will be assessed immediately postintervention as well as 16 weeks postintervention. Finally, health care use will be assessed at postintervention week 16. The control group will follow an identical outcome assessment schedule.

**Demographic and Medical Information**
Age, ethnicity, gender identity, sexual orientation, marital and parental status, years of education, income, occupation, reproductive health history, and medications will be assessed using a survey created for this study.

**Primary Outcome**
Fertility-related quality of life was assessed using the 24-item Core FertiQoL [24], yields 4 subscales: mind-body, relational, social, and emotional. High scores on the FertiQoL scale indicate a better quality of life. It is the most widely used infertility-specific measure of quality of life [25] and has been well validated in multiple studies [20]. This primary outcome was chosen in collaboration with our patient advisors as it provides an integrated measure of the emotional, physical, and interpersonal impacts of infertility.

**Secondary Outcomes**
Secondary outcomes will include depressive and anxious symptoms, instances of mood and anxiety disorders, relationship quality, and health care use. Treatment adherence and acceptability will also be assessed.

**Depressive Symptoms**
Self-reported symptoms will be assessed using the PHQ-9 [26], a 9-item measure assessing symptoms in the last 2 weeks that closely parallels the criteria for major depressive disorder as defined by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [27]. Internal consistency coefficient (ICC) has been estimated at α=0.89 and test-retest reliability at r=0.84. The PHQ-9 has been shown to be superior to other questionnaires in detecting change in anxious mood following treatment [30].

**Mood and Anxiety Disorders**

These will be assessed using NetSCID (TeleSage), the computerized version of the Structured Clinical Interview for the DSM-5 (SCID). Though the originally funded protocol proposed to use the Primary Care Evaluation of Mental Disorders (PRIME-MD), we decided to switch to the SCID because of the availability of the computerized version of this interview. The SCID is also the gold standard assessment for psychiatric diagnoses.

**Relationship Quality**
Relationship quality will be measured through the 7-item Relationship Assessment Scale. Internal consistency among individuals with infertility has been found to be high, with α=.83 [31].

**Health Care Use**
The questionnaire on health care consumption and productivity losses for patients with a psychiatric disorder will be administered 16 weeks following the intervention, asking participants to report on health care use in the last 4 months. This survey has shown good agreement with hospital- and employer-confirmed data. An additional section has been added to specifically ask about the receipt of fertility treatments.

**Treatment Acceptability**
The total number of minutes spent accessing each module will be tracked through the mobile app. Homework completion between the weekly module videos will be tracked using Qualtrics, a web-based survey platform that facilitates scheduled survey distribution, notifications, and reminders. Homework compliance will be further measured at the end of each week through the 12 items contained within the Homework Rating Scale, which assesses comprehension of homework assignments as well as effort spent on assignments [32].

**Outcomes Assessment**
A research assistant who is blinded to the participant’s treatment condition will email the participant a link to a web-based survey containing the outcome measures. If a participant fails to complete the survey within 48 hours of receipt, they will receive up to 3 reminders through email, voicemail, and SMS text messaging. Though participants will not be compensated for
completing the intervention, they will be compensated US/CAD $10 (or its equivalent depending on their location) for each postintervention survey completed (postintervention +8 biweekly follow-up surveys) and an additional US/CAD $20 for each of the 2 postintervention interviews, for a maximum total of US/CAD $130, to maximize the chances that even those participants who abandoned the intervention prematurely will complete the outcome surveys.

Statistical Analyses

Descriptive statistics will examine treatment acceptability outcomes as well as the trial recruitment rate. A 2-tailed t test will be used to compare the treatment arms in terms of baseline characteristics, assessing randomization success. A mixed model design using the MIXED procedure in SAS (version 9.4; SAS Institute) applying an intent-to-treat approach will compare the intervention arm to the waitlist control group in terms of FertiQoL, PHQ-9, GAD-7, Copenhagen Multi-Center Psychosocial Infertility Fertility Problem Stress Scale, and Relationship Assessment Scale (RAS) score across the 9 outcome measurement points (ie, at the end of intervention week 7 and biweekly for 16 weeks). Each outcome will be examined in a separate model: subject will be treated as a random effect, and the treatment assignment will be treated as a fixed effect. A repeated statement will identify assessment week as a repeated measure factor. Baseline levels of the outcome will be included as a covariate. This method has been shown to provide optimal statistical power relative to measuring pre- and postintervention outcome change [36]. In using all available data, a mixed model design has also been shown to outperform ad hoc approaches, such as the last-outcome-carried-forward approach [37].

In addition to examining the main effect of treatment assignment on outcomes, the interaction between assignment and assessment week will be examined to determine whether outcomes are maintained across the 9 postintervention measurements. Sensitivity analyses will use a similar approach to examine potential treatment moderators: membership in the LGBTQIA+ communities, baseline FertiQoL score, cultural background, disability status, and pursuit of conception through medical intervention.

The LOGISTIC procedure will assess the occurrence of mood (major or minor depressive episode or persistent depressive disorder) and/or anxiety disorders (generalized anxiety disorder, social anxiety disorder, or panic disorder) in the 2 treatment arms at the posttreatment assessment as well as at the 16-week follow-up.

To limit the family-wise error rate, the Benjamin and Hochberg [38] false discovery rate correction will be applied to all analyses.

Power Calculations

Power calculations were performed using G*Power (Axel Buchner) and are focused on the primary outcome, infertility-related quality of life. Based on SDs observed in the population of distressed women with infertility [18], setting a = .05 and power at 80%, a total of 128 participants would be needed to detect a moderate effect size (Cohen f=0.25), equivalent to a 6-point difference on the FertiQoL (out of 100) between 2 arms. To allow for a 25% (42/170) dropout rate, we will recruit 170 participants (85 per arm), in line with average completion rate of 82% observed in our meta-analysis of psychological interventions for infertility-related distress and allowing for additional dropout given considering the 16-week follow-up.

Planned Recruitment Rate

We propose to complete the trial within 2 years. The timeline relies on a recruitment rate of 3 participants per week, which we consider to be a highly conservative estimate of what is possible based on our previous experience successfully recruiting participants from this population.

Through our experience in our preliminary work, we have determined that the most successful strategy for recruiting the target population is to advertise through online infertility support or special interest groups—this will therefore be the primary method used to recruit for this study. We have found these groups to be very willing to share our research, and their members are extremely receptive as well as highly likely to be eligible to participate. We have also had great success in recruiting individuals from LGBTQIA+ communities attempting to conceive through IVF. In an ongoing study specifically targeting this population, we approached a pair of social media influencers (a lesbian couple who regularly share their experiences of undergoing IVF) who enthusiastically shared our project with their followers. Within 2 days, we had received over 400 entries in our eligibility survey.

In addition to providing large pools of highly engaged, eligible participants, one important advantage of web-based recruitment is that samples tend to be much more diverse in terms of race, education, and income relative to studies that recruit through fertility clinics, the patients for which are disproportionately high-income. While web-based recruitment can increase the risk of recruiting noneligible individuals posing as eligible participants, the use of a face-to-face Zoom enrollment session greatly reduces this risk. Our research team is also experienced in identifying suspicious survey responses.

Ethical Considerations

This study has been reviewed and approved by the University of Regina Ethics Board (REB #2023-210) as well as registered on ClinicalTrials.gov (NCT06006936). All prospective participants will provide informed written consent before enrolling in the trial.

To protect participant confidentiality, all participant data, including both interview data and questionnaire data, will be saved under ID numbers only, with no identifying information attached. Only the research team will have access to the collected survey data. The team will maintain a document associating participant names with their anonymous subject numbers. This document will be password-protected, opened only on encrypted devices, and stored separately from the rest of the data.

Participants will receive US $10 for each postintervention and follow-up survey completed and US $20 for each of the 2 postintervention interviews, for a maximum total of US $130.

https://www.researchprotocols.org/2024/1/e52662
Results

Recruitment will begin in January 2024 and continue for approximately 1.5 years. All data are expected to be collected by January 2026. Results will be uploaded on the ClinicalTrials.gov website shortly thereafter.

Discussion

Significance of the Study

It is expected that participants assigned to the Coping with Infertility program will exhibit improved fertility quality of life as well as depressive and anxious symptoms, with moderate to large effect sizes. We also expect rates of clinical mood and anxiety disorders as well as self-reported health care use to be lower among participants randomized to the treatment arm. Baseline fertility quality of life is furthermore expected to moderate the effect of treatment such that effect sizes will increase with decreasing baseline fertility quality of life. Based on the pilot study results, we expect adherence and retention to be favorable. If our hypotheses are confirmed, these findings would suggest that the Coping with Infertility program is more effective than currently available psychological interventions for infertility. Indeed, a recent meta-analysis by our team identified 58 randomized controlled trials testing psychological interventions for infertility and found that, with the exception of trials conducted in the Middle East, interventions were associated with only small psychological benefits, highlighting the need for more effective interventions [16]. The self-help nature of the Coping with Infertility program also likely makes it more cost-effective than individual psychotherapy, which typically costs US $100-$200 per session.

If our intervention proves effective, we will aim to make our mobile app publicly available through the Apple Store and Google Play Store. Decreased health care use in the treatment arm relative to the treatment-as-usual arm would provide a strong rationale for seeking government funding to upkeep the Coping with Infertility mobile app, which would allow us to make the program available free of charge. We would provide flyers and posters to fertility clinics across North America, to be posted in clinic waiting rooms and physician offices. We will reach out to relevant professional societies and nonprofit organizations, asking them to include the app as a mental health resource listed on their website. Online forums relevant to infertility will also be contacted and asked to share information related to the app. A YouTube channel will be created to house all of the weekly module videos along with a professionally produced animated explainer video introducing the intervention and describing the results from the trial supporting its efficacy. Final, we will publish our findings in open-access journal articles in respected scientific journals.

In addition to disseminating the Coping with Infertility program as a stand-alone intervention, it may also be worthwhile to pair it with other traditional mental health resources. For example, future research pairing the Coping with Infertility program with infertility support groups, or with individual psychotherapy may help target a broader audience of individuals experiencing infertility-related distress who wish to benefit from peer or therapist support. Translating the content of the program into a workbook format may also appeal to a subset of the target population.

Limitations

First, access to the Coping with Infertility program is contingent upon internet access; research participants may therefore not include individuals who do not have such access, such as those who cannot afford internet access or those living in remote communities. Second, due to the nature of the intervention, it is impossible to conduct this trial as a double-blind, randomized trial. Third, health care use will be self-reported and therefore may not capture use as accurately as hospital and clinic records.

Conclusions

This study will test a self-help program for infertility-related distress through a mobile app. If the intervention proves effective, it will provide a highly cost-effective and accessible mental health resource for those struggling with the mental health impacts of infertility. This will fill an important gap in the need for more effective interventions [16]. The self-help nature of the Coping with Infertility program also likely makes it more cost-effective than individual psychotherapy, which typically costs US $100-$200 per session.

Acknowledgments

The authors would like to thank Canadian Institutes of Health Project Grant (#PJT186221) for the funding to complete this trial.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Peer review report from the Canadian Institutes of Health Research.

[PDF File (Adobe PDF File), 80 KB - resprot_v13i1e52662_app1.pdf ]

References


3. Stanger S. Exploring the Experiences of Members of 2SLGBTQIA+ Communities Pursuing Third-Party Reproduction [Honours Thesis]. Regina, Saskatchewan: University of Regina; 2022. URL: https://ourspace.uregina.ca/items/bd3716b4-3966-4df3-a79f-9f8e6ee66e99 [accessed 2023-12-08]


Abbreviations

CONSORT-SPI: Consolidated Standards of Reporting Trials for Social and Psychological Intervention Trials
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
FertiQoL: Fertility Quality of Life Scale
GAD-7: Generalized Anxiety Disorder-7
ICC: internal consistency coefficient
IVF: in vitro fertilization
LGBTQIA+: lesbian, gay, bisexual, transgender, queer, intersex, and asexual
PHQ-9: 9-item Patient Health Questionnaire
PRIME-MD: Primary Care Evaluation of Mental Disorders
SCID: Structured Clinical Interview for the DSM-5
Protocol

Effects of a Powered Ankle-Foot Prosthesis and Physical Therapy on Function for Individuals With Transfemoral Limb Loss: Rationale, Design, and Protocol for a Multisite Clinical Trial

Jason T Maikos¹*, PhD; Alison L Pruziner²,³,⁴*, PT, DPT; Brad D Hendershot²,³,⁵*, PhD; David V Herlihy⁶*, BS; John M Chomack⁶, MS; Michael J Hyre⁶*, MS, CAPM; Samuel L Phillips⁷*, PhD, CP; Alexis N Sidiropoulos¹*, PhD; Christopher L Dearth²,³,⁵*, PhD; Leif M Nelson¹,³,⁴*, PT, DPT

¹Veterans Affairs New York Harbor Healthcare System, New York, NY, United States
²Walter Reed National Military Medical Center, Bethesda, MD, United States
³Extremity Trauma and Amputation Center of Excellence, Falls Church, VA, United States
⁴National Veterans Sports Programs and Special Events, Department of Veterans Affairs, Washington, DC, United States
⁵Uniformed Services University of the Health Sciences, Bethesda, MD, United States
⁶The Narrows Institute for Biomedical Research and Education, Brooklyn, NY, United States
⁷James A. Haley Veterans' Hospital, Tampa, FL, United States
*
all authors contributed equally

Corresponding Author:
Jason T Maikos, PhD
Veterans Affairs New York Harbor Healthcare System
423 E. 23rd Street
New York, NY, 10010
United States
Phone: 1 2126867500 ext 7482
Email: jason.maikos@va.gov

Abstract

Background: Powered ankle-foot prosthetic devices can generate net positive mechanical work during gait, which mimics the physiological ankle. However, gait deviations can persist in individuals with transfemoral limb loss because of habit or lack of rehabilitation. Prosthetic research efforts favor the design or evaluation of prosthetic componentry and rarely incorporate any type of rehabilitation, despite evidence suggesting that it is critical for minimizing gait imbalances. Given the accelerated rate of innovation in prosthetics, there is a fundamental knowledge gap concerning how individuals with transfemoral limb loss should learn to correctly use powered ankle-foot devices for maximum functional benefit. Because of the recent advances in prosthetic technology, there is also a critical unmet need to develop guidelines for the prescription of advanced prosthetic devices that incorporate both physical and psychological components to identify appropriate candidates for advanced technology.

Objective: The primary goal of this investigation is to examine the roles of advanced prosthetic technology and a device-specific rehabilitative intervention on gait biomechanics, functional efficacy, and pain in individuals with transfemoral limb loss. The secondary goal is to develop preliminary rehabilitation guidelines for advanced lower limb prosthetic devices to minimize gait imbalances and maximize function and to establish preliminary guidelines for powered ankle-foot prosthetic prescription.

Methods: This prospective, multisite study will enroll 30 individuals with unilateral transfemoral limb loss. At baseline, participants will undergo a full gait analysis and assessment of function, neurocognition, cognitive load, subjective preferences, and pain using their current passive prosthesis. The participants will then be fitted with a powered ankle-foot device and randomized into 2 equal groups: a powered device with a device-specific rehabilitation intervention (group A) or a powered device with the current standard of practice (group B). Group A will undergo 4 weeks of device-specific rehabilitation. Group B will receive the current standard of practice, which includes basic device education but no further device-specific rehabilitation. Data collection procedures will then be repeated after 4 weeks and 8 weeks of powered ankle use.

Results: This study was funded in September 2017. Enrollment began in September 2018. Data collection will conclude by March 2024. The initial dissemination of results is expected in August 2024.

Conclusions: The projected trends indicate that the number of individuals with limb loss will dramatically increase in the United States. The absence of effective, evidence-based interventions may make individuals with transfemoral limb loss more susceptible...
to increased secondary physical conditions and degenerative changes. With this expected growth, considerable resources will be required for prosthetic and rehabilitation services. Identifying potential mechanisms for correcting gait asymmetries, either through advanced prosthetic technology or rehabilitative interventions, can provide a benchmark for understanding the optimal treatment strategies for individuals with transfemoral limb loss.

**Trial Registration:** ClinicalTrials.gov NCT03625921; https://clinicaltrials.gov/study/NCT03625921

**International Registered Report Identifier (IRRID):** DERR1-10.2196/53412

---

**Introduction**

**Background**

There are approximately 1.9 million Americans with limb loss today, with an estimated 185,000 people who undergo an amputation procedure each year [1]. Over the last 2 decades, the Department of Veterans Affairs (VA) and the Department of Defense (DoD) have experienced an increase in the number of veterans and service members with lower limb loss [2]. Since the start of the most recent conflicts, more than 1700 service members have experienced combat-related limb loss, with the vast majority of these traumatic amputations of the lower limb [3-4]. VA, the largest integrated health care network in the United States, serves this unique population after their separation from active duty and provides care for an additional 41,000 veterans with lower limb loss [5]. Transfemoral limb loss, the second most common level of lower limb loss, accounts for one-fifth of the total limb loss population in the United States [6]. With this already large population expected to grow, effective outcomes-based clinical practice will be necessary to improve mobility, decrease long-term disability, and provide a higher quality of life.

**Abnormal Gait Mechanics for Individuals With Transfemoral Limb Loss**

Individuals with transfemoral limb loss have unique functional challenges owing to the loss of the knee and ankle joints [7-9]. Gait mechanics of individuals with transfemoral limb loss have been extensively investigated, with abnormalities typically characterized by asymmetries in stance phase biomechanics [10-13]. Individuals with transfemoral limb loss exhibit increased ipsilateral hip extensor activity and hip power, which is thought to be compensation for the lack of ankle power normally provided by the gastroc-soleus complex [14]. Consequently, compensatory mechanisms at joints proximal to the level of limb loss are often used to replace the function normally delivered by the muscles surrounding the ankle joint [12]. Individuals with unilateral transfemoral limb loss also tend to walk with longer stance times on the intact versus prosthetic limb [10], which can lead to a corresponding asymmetrical load distribution [15]. These asymmetrical joint forces place greater demands on the intact limb, which may explain the higher prevalence of musculoskeletal injuries, pain, and joint degeneration of the intact limb compared with uninjured individuals [16,17]. Significant among these secondary conditions is pain, specifically in the intact knee and lower back.

In a study of experienced prosthesis users, knee pain in the intact limb was the primary complaint of 75% of individuals with transfemoral limb loss [18]. In a sample of 63 male veterans with traumatic lower limb loss, individuals with transfemoral limb loss were 5 times more likely to have intact knee pain compared with neurotypical participants [16]. Among individuals with lower limb loss (both transtibial and transfemoral), 71% reported back pain within the previous month, but individuals with transfemoral versus transtibial limb loss were significantly more likely to have greater pain intensity [19]. Chronic, persistent pain can lead to limitations in function. There is a significant need to explore the effects of advanced prosthetic technologies and rehabilitative interventions on pain reduction, function, and biomechanics.

**Biomimetic Prosthetic Technology**

New technologies in lower limb prostheses have attempted to combat gait pathologies by generating biomimetic ankle power through spring-clutch mechanisms or advanced sensor and actuator technology [20]. Recent advances in microelectronics, battery technologies, and the development of several new types of actuators [21,22] have ushered in the development of powered lower limb prostheses that can better replicate the positive work phases of the ankle through the use of actuators, motors, or pneumatic muscles [23-27].

The Empower (OttoBock Inc), which uses a series-elastic actuator and a carbon Composite footplate, is currently the only commercially available powered ankle-foot device [28,29]. The Empower has been investigated in the population of individuals with transtibial limb loss, but it has yet to be fully investigated in individuals with transfemoral limb loss [30-33]. In a study of individuals with transtibial limb loss, the use of a powered versus passive ankle-foot device reduced the peak resultant force and knee adduction moment on the unaffected leg during level ground walking, potentially limiting the risk of secondary musculoskeletal comorbidities [31]. Individuals with transtibial limb loss using the same powered ankle-foot device had improved ankle power, greater net trailing limb step-to-step transition work, and a lower metabolic rate compared with a passive energy storing and returning ankle-foot prosthesis during level ground ambulation [32].

Although biomimetic prosthetic devices can better approximate biological ankle biomechanics, residual gait deviations can persist, either because of habit or a lack of proper rehabilitative training. For example, despite greater ankle power generation with powered ankle-foot device use, individuals with transtibial...
limb loss can still walk with compensatory strategies at the proximal joints, which can be attributed to the introduction of new interlimb asymmetries from the uniarticular function of the device [30]. Therefore, device-specific rehabilitation may be needed to minimize or correct the reported deficiencies. Similarly, in the absence of an evidence-based rehabilitation program to correct or minimize preexisting gait asymmetries, instrumented gait analyses that assess the biomechanical function of prosthetic devices may be more likely to quantify the physical gait deviations developed through habit or lack of training rather than device-specific attributes [34]. Therefore, it may be more accurate to postulate that powered ankle-foot devices, through the generation of normative ankle power during push off, offer an opportunity to minimize gait deviations and normalize prosthetic function but not without the incorporation of a rehabilitation program to train prosthesis users to reduce existing gait deviations.

Prosthetic Rehabilitation Programs

The current state of prosthetic research efforts appears to favor the design and evaluation of prosthetic componentry, particularly with respect to gait mechanics, and rarely incorporates or reports any type of physical therapy (PT) program or device-specific training [34]. Given the accelerated rate of technological innovation in prosthetic devices, there is a fundamental knowledge gap concerning how individuals with lower limb loss should learn to correctly use this advanced, powered technology for maximum benefit. However, previous investigations have examined the effectiveness of rehabilitation protocols on the outcomes of individuals with transfemoral limb loss who used passive prosthetic devices. Prosthetic gait training based on proprioceptive feedback for individuals with transfemoral limb loss was more effective for improved weight-bearing and temporal-spatial parameters than traditional gait training [35]. Sjodahl et al [36] used instrumented gait analysis to measure the gait parameters of individuals with unilateral transfemoral limb loss before and after a training program and reported improved walking speed and sagittal plane hip kinematic symmetry after training. However, the authors also reported increases in compensatory strategies for the intact limb, including an increase in the intact knee extension moment. Virtual reality–based gait training with real-time biomechanical feedback improved frontal plane hip, pelvis, and trunk motion during level ground walking [37]. Currently, there have been no published studies detailing the effects of a device-specific rehabilitation program on the biomechanical or functional outcomes of individuals with transfemoral limb loss who use a powered ankle-foot prosthesis. In this investigation, this knowledge gap will be addressed, and a benchmark to understand optimal treatment strategies will be provided for individuals with transfemoral limb loss to minimize gait impairments.

Summary

The development of evidence-based health care practices is critical to maximizing prosthetic and health outcomes in the growing population of individuals with transfemoral limb loss. Identifying potential mechanisms for correcting gait asymmetries, through advanced prosthetic technology, rehabilitative interventions, or both, can provide a benchmark to better understand the optimal treatment strategies for individuals with transfemoral limb loss. Despite research suggesting that an evidence-based rehabilitation program that incorporates prosthetic gait training is a critical factor in minimizing compensatory mechanisms [38-40], most prosthetic device protocols fail to incorporate any type of significant rehabilitation or device-specific training. Therefore, this investigation will be the first to elucidate the effects of an advanced powered prosthesis and the role of rehabilitative interventions on gait biomechanics, performance, and pain in individuals with transfemoral limb loss.

Study Objectives

The overarching goal of this investigation is to examine the roles of advanced prosthetic technology and a device-specific rehabilitative intervention in individuals with transfemoral limb loss. The central hypothesis is that powered plantarflexion, coupled with an evidence-based, device-specific PT intervention, will improve biomechanical outcomes, which will correlate with decreased pain and improved functional performance. The objectives of this investigation are as follows:

1. To examine the biomechanical and functional efficacy of a powered prosthesis compared with a passive prosthesis for individuals with transfemoral limb loss
2. To determine the effects of a powered prosthetic ankle-foot device and a PT intervention on lower extremity kinematic and kinetic patterns, functional efficacy, and pain in individuals with transfemoral limb loss
3. To develop preliminary rehabilitation guidelines for a powered ankle-foot device to minimize gait imbalances and maximize function, as well as to establish preliminary guidelines for powered ankle-foot prosthetic prescription

Methods

Study Overview

This investigation will be a prospective, multisite study including VA New York Harbor Healthcare System (VANYHHS), James A. Haley Veterans’ Hospital (JAHVH), and Walter Reed National Military Medical Center (WRNMMC). Enrollment began in September 2018, and data collection is expected to conclude in 2024. Briefly, 30 individuals with transfemoral limb loss are expected to be enrolled equally across the 3 sites. For all participants, a full biomechanical gait analysis, functional measures, surveys, neurocognitive assessment, cognitive load assessment, and pain assessment will be captured at baseline with their clinically prescribed passive energy storing and returning ankle-foot prosthesis. The participants will be fitted with a powered ankle-foot device (Empower) and then be evaluated for safe use. The participants will then be randomly assigned into 2 groups: a powered ankle-foot device with a 4-week, 8-session device-specific PT intervention (group A) or a powered ankle-foot device with the current standard of practice (group B), which includes basic device education and training provided by the study prosthetist (outlined in the Powered Ankle-Foot Device Standard of Practice section), but no device-specific PT intervention. Group A will then undergo 4 weeks of
device-specific rehabilitation, while group B will not receive any further PT. All participants will then undergo a full gait analysis as well as assessments of function, subjective preferences, neurocognition, cognitive load, and pain after 4 weeks and 8 weeks of powered ankle-foot device use (Figure 1). A comparison between the 2 groups will help evaluate the efficacy of a powered versus passive prosthesis, as well as elucidate the contribution of device-specific effects to rehabilitation-specific effects for individuals with transfemoral limb loss.

Figure 1. Participant timeline of activities.

Participants
A convenience sample of 30 individuals with unilateral transfemoral limb loss will be recruited for this study (Textbox 1 shows the inclusion and exclusion criteria). All participants will consent to participate before participating in any study activities. The participants will be randomly stratified into 2 study arms: group A, a powered ankle-foot device with device-specific PT (15/30, 50%), and group B, a powered ankle-foot device with standard of practice (15/30, 50%), which includes basic device education and training but no device-specific PT intervention. Recruitment of participants will be on a first-come, first-serve basis among the patients of VANYHHS, JAHVH, and WRNMMC. The participants will be recruited through the VANYHHS, JAHVH, and WRNMMC rehabilitation and prosthetic clinics. All participants will have experience using a microprocessor knee and will currently use a passive-elastic ankle-foot prosthesis.
Textbox 1. Inclusion and exclusion criteria.

### Inclusion criteria
- Unilateral transfemoral limb loss because of any etiology
- Use of a microprocessor knee with >6 months of experience
- ≤8 limb loss–related physical therapy sessions in the previous 6 months
- Aged at least 18 years
- Score of ≥33 on the Amputee Mobility Predictor, corresponding to a high K2 or above ambulator
- Able to walk a minimum of 30 meters without an assistive device
- Able to walk on a treadmill for 5 minutes at self-selected speed with or without the use of handrails

### Exclusion criteria
- Inability to tolerate the wearing of a socket or a poorly fitting socket
- Conditions of the intact limb prohibit prosthesis use (eg, ulcers, sores, skin breakdown, burns, poor skin coverage, contractures, and severe heterotopic ossification)
- The length of the residual limb prohibits socket or prosthesis fitting
- Cognitive deficits or a mental health pathology limiting the ability to participate fully in the study or any deficit deemed by the principal investigator to be detrimental to the completion of the study
- Significant comorbidity, which would interfere with the study (eg, neuropathy, uncontrolled diabetes, receiving dialysis, insensate feet, severe phantom pain, or a history of skin ulcers)
- Severe circulatory problems, including peripheral vascular disease and pitting edema
- Pregnant women in the second trimester or beyond or women who will be in the second trimester within the enrollment period
- Weigh >130 kg at screening
- Use of nonprescribed opioids or overuse of any prescription drugs
- Major upper limb loss
- Currently uses a powered ankle-foot prosthesis as a primary prosthesis or used a powered ankle-foot device as a primary prosthesis in the previous 6 months
- Any cardiopulmonary, metabolic, or integumentary diagnosis where walking for 15 minutes is contraindicated

### Ethical Considerations
This study was approved by the following institutional review boards (IRBs): VANYHHS IRB (1643), WRNMCC IRB (WRNMCC-M-2018-0167), and the University of South Florida IRB, which is the IRB of record for JAHVH (IRB STUDY000870). The oversight and protection of human participants was also approved by the US Army Medical Research and Development Command Office of Human Research Oversight (EO4081). All participants will provide informed consent before participating in any study activities.

### Baseline Visit

#### Informed Consent, Enrollment, and Randomization

Informed consent for each potential participant will be conducted in person in a private room. The site-principal investigator or qualified designee will explain the study protocol in detail. The participants will be asked to consent to randomization of the treatment group, either to be fit with the powered ankle-foot device and receive device-specific PT (group A) or to be fit with the powered ankle-foot device and receive the current standard of practice that does not include device-specific PT (group B). The participants will be asked to make a commitment to be available for all study-related activities. The individuals will be given adequate time to review and comprehend all information about the study before agreeing to participate, minimizing the possibility of coercion and undue influence. After the study has been explained and consent has been given, the participants will be randomized into the 2 groups using a computer-generated algorithm that block randomizes participants into each group at each site.

#### Baseline Data Capture

Regardless of the group assignment, all participants will undergo baseline data capture using their current passive energy storing and returning prosthesis. The study prosthetist will ensure proper fit and alignment of the prosthesis before any data collection. Once fit and alignment are confirmed, the participants will complete 3 surveys: an assessment of quality of life, the Prosthesis Evaluation Questionnaire (PEQ), and the PEQ Addendum. Then, pain, cognitive load, neurocognition, biomechanical gait analysis, and functional outcomes will be captured at baseline for all participants on their current passive energy storing and returning prosthesis.
Quality-of-Life Assessment

A single-item assessment will ask the participants to rate their quality of life over the past 4 weeks on a 100-mm visual analog scale.

Assessment With PEQ

The PEQ is a self-reported visual analog–style questionnaire for people with lower limb loss who use a prosthesis to evaluate the prosthesis and life with the prosthesis. The PEQ is organized into 9 domains that may be used independently to measure a specific domain of interest. Domains of utility, appearance, sounds, residual limb health, and ambulation will be used in this investigation. In addition, the PEQ contains items beyond the domains that can be evaluated individually, including questions on satisfaction, pain, and transfers that will be used in this study [41].

Assessment With PEQ Addendum

The PEQ Addendum asks 2 open-ended questions assessing any falls or stumbles that the participant may have experienced over the previous 4 weeks [42].

Pain Assessment

Participants will complete the Patient-Reported Outcomes Measurement Information System Pain Interference Scale 8a, which is a self-report survey that assesses the extent to which pain interferes with physical, psychosocial, cognitive, emotional, and recreational activities [43].

Neurocognitive and Cognitive Load for Prosthesis Use

Prosthesis use requires physical capabilities and the cognitive capacity to learn new techniques across different situations and environments. These skills include spatial organization, memory, attention, and visuospatial function [44]. Powered ankle-foot devices can be more complex than passive devices and may require certain levels of neurocognition and cognitive load for ambulation, especially for higher-level functional tasks. Diabetes and peripheral vascular disease, the most prevalent causes of lower limb loss, are linked to declining neurocognition [45,46]. Importantly, diminished neurocognitive function is not often observed until late in the rehabilitation process [well after prosthetic prescription] [47], which can result in the mismanaged use of staff and patient time and prosthetic resources. If certain levels of neurocognitive abilities correlate with successful prosthetic outcomes with the powered prosthesis, neurocognitive assessment (before prosthetic prescription) can potentially aid in the selection of appropriate candidates for advanced technology.

Cognitive Load Assessment

Cognitive load assessments will be performed at baseline using the passive energy storing and returning ankle-foot prosthesis. Before any cognitive load testing, a self-selected walking speed will be established via treadmill walking. The participants will be encouraged to walk unsupported on the treadmill, if possible. The treadmill console will be covered to prevent number distractions, and the participants will be reminded to hold their gaze straight ahead. To determine the self-selected speed, the participants will begin walking on the treadmill at a comfortable speed in the absence of any additional cognitive load. The treadmill will then be increased by 0.09 m/s every 10 seconds until the participant verbalizes their preferred speed. To avoid quick trigger responses, the speed will be increased by 0.18 m/s and then subsequently lowered by 0.04 m/s every 5 seconds until the participant verbalizes their preferred speed. If the final speed does not match the initial speed, this procedure will be repeated until the participant matches the preferred walking speed.

Cognitive load testing consists of five 1-minute standardized cognitive tasks: 1 serial subtraction task, 3 controlled oral word association tasks, and 1 category task. These tasks consist of auditory and verbal cognitive measures to simulate real-world conditions. Visual tasks will not be used to avoid measures that require reading while walking. The participants will complete each cognitive task while walking on a treadmill at the previously noted self-selected speed. The directions will be read to each participant to ensure protocol consistency. The number of correct answers will be determined for each 1-minute test using a digital voice recorder and paper recordings to ensure accurate documentation of the participants’ answers. The participants will complete a practice trial before each task initiation to ensure complete comprehension. The participants will also be allowed to rest between cognitive tests as needed. The cognitive tasks are as follows:

- **Serial subtraction**: Serial subtraction is a mental arithmetic task [48]. The participants are given a random 3-digit number and asked to continually subtract 3 while they walk for 1 minute. The number of errors will be calculated. The participants are not penalized for multiple errors if 1 error was made; however, they continued sequentially thereafter.
- **Controlled Oral Word Association Test**: This measure consists of 3 tests that measure verbal phonemic fluency and other neuropsychologic domains [49]. The participants will be asked to list words beginning with a certain letter while walking for 1 minute. The test is then repeated for 2 other letters. The total number of unique words for each letter will be documented.
- **Category Test**: The participants are asked to list words belonging to a certain category within 1 minute (eg, fruits or parts of a car). Category naming has shown validity and reliability [50]. The total number of unique words for each category will be documented.

After each cognitive load test, the participants will be asked to rate on a 0 to 10 scale (with 0 being “none” and 10 being “a great deal”) their focus on walking, their focus on thinking of words or subtracting numbers, and their focus on distractions. This will provide information on the subjective experience of cognitive load.

Neurocognitive Assessment

Following the cognitive load assessment, the participants will take an electronic neurocognitive battery (CNS Vital Signs) [51]. The computerized neurocognitive assessment measures 5 domains (memory, psychomotor speed, reaction time, complex attention, and cognitive flexibility), is designed to be administered serially, and has demonstrated good test-retest reliability. The computerized neurocognitive battery measures the participant's ability to perform a variety of tasks that require attention, memory, and cognitive flexibility. The tasks include:

- **Category Test**: The participants are asked to list words belonging to a certain category within 1 minute. The number of unique words for each category will be documented.
- **Controlled Oral Word Association Test**: This measure consists of 3 tests that measure verbal phonemic fluency and other neuropsychologic domains. The participants will be asked to list words beginning with a certain letter while walking for 1 minute. The test is then repeated for 2 other letters. The total number of unique words for each letter will be documented.
- **Serial subtraction**: Serial subtraction is a mental arithmetic task. The participants are given a random 3-digit number and asked to continually subtract 3 while they walk for 1 minute. The number of errors will be calculated. The participants are not penalized for multiple errors if 1 error was made; however, they continued sequentially thereafter.
- **Category Test**: The participants are asked to list words belonging to a certain category within 1 minute. The number of unique words for each letter will be documented.
reliability. The neurocognitive assessment consists of the following 7 tests:

- Verbal memory: the participants are instructed to remember 15 words that are displayed 1 at a time every 2 seconds. The target words are then randomly mixed with 15 new words. The participants are instructed to press the space bar when a target word is displayed. This test is repeated at the end of the assessment with the same 15 target words randomly mixed with 15 new nontarget words.
- Visual memory: the participants are instructed to remember 15 geometric shapes that are shown 1 at a time every 2 seconds. The shapes are then randomly mixed with 15 new nontarget shapes. The participants are instructed to press the space bar when they identify a target shape. This test is repeated at the end of the neurocognitive assessment period.
- Finger tapping: the participants are instructed to tap the space bar with their right index finger as quickly as possible during the 10-second test. The test is then repeated with the left index finger.
- Symbol Digit Coding: this test consists of 8 digit-symbol pairs, followed by a list of digits. The participants are instructed to serially type numbers that correspond to the symbols during the 120-second test.
- Stroop test: in part 1 of the Stroop test, the participants are randomly shown the words “GREEN,” “YELLOW,” “RED,” and “BLUE.” These words are printed in black. The participants are instructed to press the space bar when they see one of these words. In part 2, the same words appear on the screen but are printed in different colors. The participant is instructed to press the space bar when the color presented matches the word (eg, the word “RED” is printed in red). In part 3, the same words are presented on the screen in different colors. The participants are then instructed to press the space bar when the color presented does not match the word (eg, the word “RED” is printed in green).
- Shifting Attention: in this test, 3 figures appear on the screen. There is a single figure at the top of the screen (either a circle or a square). At the bottom of the screen, 2 figures are presented (both a square and a circle). The figures are randomly mixed to be either red or blue. The participant is then instructed to match one of the corresponding bottom figures with the top figure by shape or color. However, the rules for matching by shape or color change at random. This test occurs for 90 seconds.
- Continuous performance test: during a 5-minute period, the participants are asked to press the spacebar when the letter B appears on the screen. The participants are further instructed not to respond to any other letters. The letters are presented at random.

**Biomechanical Gait Analysis**

Gait analyses will be performed at the biomechanics laboratories at VANYHHS, WRNNMC, and JAHVH. The VANYHHS laboratory is a 133-m^2 space comprising an 11-camera motion capture system (Qualisys, Inc) with 4 multiaxis force platforms (AMTI Inc). At WRNNMC, the biomechanics laboratory is an 167-m^2 space comprising an 18-camera motion capture system (Qualisys, Inc) and 6 multiaxis force platforms (AMTI Inc). At JAHVH, the motion capture laboratory is a 74-m^2 space equipped with a 12-camera motion capture system (Vicon Inc) and 4 multiaxis force platforms (AMTI Inc). All 3 systems track the positions of passive reflective markers at a rate of 120 Hz, and force platforms sample ground reaction forces at a rate of 1200 Hz. Visual3D software (C-Motion Inc) will be used for the analysis of 3D motion capture data.

All laboratories in this investigation will follow recommendations provided by an interlab reliability study that was conducted between gait laboratories at the 3 military treatment facilities [52]. Specifically, all sites will use identical marker sets, identical anatomical segment definitions, and a single examiner at each site to conduct postprocessing of the respective data to reduce potential variability between the laboratories.

For all participants, biomechanical gait analysis will be performed at baseline using their prescribed energy storing and returning prostheses. All kinematic and kinetic biomechanical measures will be captured using an identical 6-degrees-of-freedom marker set. A custom, full-body passive reflective marker set will be placed on each participant, which tracks each segment independently, allowing for the accurate measurement of movements. As previously described [53], 78 markers will be placed or digitized on the head, trunk, pelvis, and extremities. Marker placements for the prosthetic limb will be matched to those of the intact leg or placed on the centers of rotation of the prosthetic ankle-foot and knee devices. The cluster technique will be used to minimize the surface-to-bone displacements for the thigh, shank, and upper arm–mounted markers [54]. Tracking clusters will be placed bilaterally on the thigh, the tibial crest, and the upper arm. Functional joints, adapted from Schwartz and Rozumalski [55], will also be calculated for the intact ankle and knee as well as bilaterally for the hips.

During each experimental session, the participants will separately walk at 3 speeds across an instrumented walkway until 5 acceptable trials for each foot at each speed are completed. Trials will be considered acceptable when a foot makes full contact with a force platform. Because kinetic outcome measures are speed-dependent, the participants will ambulate at 3 controlled speeds: 0.7, 1.0, and 1.3 m/s. These speeds were selected to represent a slow, moderate, and fast walking speed, respectively, for individuals with transfemoral limb loss. The order of speeds will be randomized for each data collection visit. Auditory feedback will be provided to the participant by the study team to ensure that all participants walk at the targeted speed (−5% to +5%). The main purpose of this session is to collect joint motion, force, torque, and power data at each walking speed. The ranges of motion; speeds and accelerations; and hip, knee, and ankle joint moments of force and generated and absorbed powers will be computed using inverse dynamics methods. Temporal-spatial parameters will also be recorded.

The reflective marker positions will be digitized using motion tracking software. A 15-segment rigid body model (head, trunk,
pelvis, bilateral upper and lower arms, hands, thighs, shanks, and feet) will be created based on the skin-mounted markers and functional joints. Local coordinate systems for each segment will be defined using the International Society of Biomechanics recommendations [56,57]. The data of 5 acceptable walking trials at each speed will be processed using Visual3D. Marker data will be filtered with a 6 Hz Butterworth low-pass filter. Raw analog data will be filtered using a second-order low-pass Butterworth filter with a 25-Hz cutoff frequency. Visual3D will be used to calculate temporal-spatial values, walking speed, and lower extremity kinematics and kinetics. Inverse dynamic analysis will be applied to the kinematics of the biomechanical model and to the location, magnitude, and direction of ground reaction forces acting on the foot to calculate lower extremity joint torques and powers, including ankle, knee, and hip power of the biological and prosthetic limb over the stance phase, as well as the frontal plane knee moments for the unaffected leg.

**Functional Outcome Measures**

The effects of the prosthetic devices and the rehabilitative intervention on physical performance will be evaluated using agility and mobility tests, including the 6-minute walk test (6MWT) [58], the Amputee Mobility Predictor with prosthesis (AMPpro) [59], and the Comprehensive High-Level Activity Mobility Predictor (CHAMP) [60]. By capturing the functional measures in each group, the effects of the ankle-foot device can be isolated from the rehabilitation effects on physical performance. These measures are as follows:

- **6MWT**: the 6MWT measures the distance an individual can walk in 6 minutes without help or encouragement. It is a valid and reliable measure that correlates with physical function and has good interrater and intrarater reliability in individuals with lower limb loss [58].
- **AMPpro**: the AMPpro is a 21-item instrument designed to measure prosthetic mobility in individuals with lower limb loss [59].
- **CHAMP**: participants who attain a score of 37 or higher on the AMPpro will undergo the CHAMP, which consists of the following tasks:
  - **Single limb stance**: participants fold their arms across their chest and then lift their foot above a 15-cm cone or box. The test ends when the foot touches the ground again (or until 30 s) or if the arms uncross. This procedure is performed on both feet.
  - **Edgren Sidestep Test**: participants sidestep left and then right along a 5-meter line of cones (1 meter apart). The sidestep test lasts for 10 seconds.
  - **T-Test**: the T-Test measures forward, lateral, and backward walking (or running) and sidestepping in a T pattern.
  - **Illinois Agility Test**: this advanced agility test requires the participant to run or walk and change direction around multiple cones. Over 60 meters, participants perform 90° and 180° turns 11 times around multiple cones.

---

**Powered Ankle-Foot Prosthesis Fitting**

Following all baseline data collection procedures, all participants will then be fit with the Empower. Study prosthetists at VANYHHS, WRNMMC, and JAHVH are highly experienced in fitting all commercial microprocessor knees and the Empower. The study prosthetists will bench align the powered ankle-foot device onto the participant’s existing microprocessor prosthetic knee and socket. Once the powered ankle-foot device is fitted and bench-aligned to the prosthetic knee and socket, dynamic alignment of the prosthetic knee and ankle will occur. Initially, the study prosthetist will ensure proper alignment of the microprocessor prosthetic knee with the Empower turned off. The Empower will still function (ie, articulate) with the power off but will not provide net positive plantarflexion torque. The microprocessor prosthetic knee software and prosthesis alignment will be adjusted during standing and walking tasks, as necessary, until the prosthetist and the participant are satisfied with the knee alignment. Once the microprocessor knee setup is completed, the Empower will be powered on and adjusted according to the manufacturer’s specifications. Briefly, the participants will sequentially walk at 3 different speeds (slow, self-selected, and fast) while the stiffness and power delivery of the powered ankle-foot device is tuned [33]. If further dynamic adjustments to prosthetic knee alignment are necessary, these adjustments will be made at this point.

**Powered Ankle-Foot Device Standard of Practice**

Once a stable and comfortable alignment has been established, all participants will be educated by the study prosthetist on the proper use of the Empower, which includes battery handling and charging, understanding low battery indicators, considerations while driving with the Empower (if applicable), and avoidance of exposure to rain and water. Next, the participants will ascend and descend an Americans with Disabilities Act–compliant ramp under the supervision of the study physical therapist or prosthetist. The intent of ramp walking is to trigger the power for ascent but not during descent. Each participant will be given the opportunity to practice the ramp as often as necessary to ensure safe and comfortable use. The participants will then negotiate a standard staircase with handrails under the supervision of the study physical therapist or prosthetist to ensure that they can safely negotiate stairs using the Empower. Proper technique will consist of demonstrating the correct foot placement on each step to activate powered push off during ascent. Participants unable to ascend stairs in a step-over-step pattern will be shown the correct foot placement using a step-to gait pattern. For proper stair descent, participants will be shown the correct foot placement to initiate rollover and not trigger the power. Participants will be given time to practice stair ascent and descent. After stair and ramp ambulation, participants will be asked to demonstrate safe use of the Empower in different situations, including turning, varying speeds, sudden stops, obstacle avoidance, stepping over obstacles, and different surfaces. Once the physical therapist is satisfied that the participant has demonstrated safe use and all questions have been answered, the participant will be released home with the Empower. If the physical therapist feels that the progress is unsatisfactory, the participant will not take the Empower home and will be asked to return for continued
supervised use until the participant demonstrates safe use. The standard of practice to use the powered prosthesis is approximately 30 to 45 minutes after fitting and tuning.

Following the baseline visit, group A will undergo the device-specific PT program, whereas group B will not undergo any further training.

PT Program: Group A

Overview

Participants in group A will complete, on average, 8 PT sessions lasting 30 to 45 minutes each. The exact PT protocol and criteria for advancement are outlined by level in the subsequent section. In brief, level 1 of the PT plan will focus on education, strengthening through therapeutic exercises, and early neuromuscular reeducation. A home exercise program (HEP) will be initiated during the initial sessions and will progress along with the program. Level 2 will include gait training on level surfaces, sit-to-stand transitions, and ramp negotiation. Level 3 will include multidirectional training for both neuromuscular reeducation and gait and the introduction of stair ascent and descent. Training will conclude with level 4 where the previous skills will be further challenged and advanced gait skills, including ambulation on ramps and uneven surfaces, will be introduced. Participants must meet the outlined criteria before progressing through each level. Participants who do not meet the specified criteria will be offered additional PT sessions. The number of additional sessions will be recorded and used to refine the PT program for future use.

Level 1 (Sessions 1 and 2)

Level 1 (Table 1) includes initial evaluation, patient education, gait assessment, training to ensure safe use in the community, therapeutic exercises (including introduction of the HEP), and the initiation of early neuromuscular reeducation training.

Table 1. Level 1 device-specific physical therapy protocol.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Description</th>
<th>Criteria for advancement to level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>Assessment of prosthesis fit and gait to determine deficits</td>
<td>N/A</td>
</tr>
<tr>
<td>Strengthening</td>
<td>Strengthening of transversus abdominis and multifidus, gluteus maximus,</td>
<td>Within normal limits for range of motion</td>
</tr>
<tr>
<td></td>
<td>gluteus medius, and general trunk strengthening</td>
<td></td>
</tr>
<tr>
<td>Stretching</td>
<td>Address deficits from evaluation, including iliopsoas</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Explanation of function of powered ankle-foot device</td>
<td></td>
</tr>
<tr>
<td>Gait training</td>
<td>Safely negotiate level surface without increased falls or stumbles</td>
<td></td>
</tr>
<tr>
<td>Mobility training</td>
<td>Side stepping, backward stepping, and turns</td>
<td></td>
</tr>
<tr>
<td>Neuromuscular reeducation</td>
<td>Weight shifting and control over prosthesis, intact limb mobility</td>
<td>Independent in HEP</td>
</tr>
<tr>
<td></td>
<td>(toes in and toes out, heel in and heel out) to promote weight shifting,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>static single limb balance training on prosthesis side with upper limb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>support, and anterior and posterior stepping exercises with the intact limb</td>
<td></td>
</tr>
<tr>
<td>HEP</td>
<td>Initiated with all therapeutic exercises outlined in level 1</td>
<td>Independent in HEP</td>
</tr>
</tbody>
</table>

Table 1a: not applicable.

bHEP: home exercise program.

Level 2 (Sessions 3 and 4)

Level 2 (Table 2) will include the progression of therapeutic exercises through increased frequency, duration, and resistance. All strengthening and stretching will be shifted to the HEP by the completion of level 2. Lumbar, abdominal, and closed kinetic chain lower extremity strengthening exercises will progress to more dynamic positions. Single limb stance progressions (neuromuscular reeducation) will include decreased upper limb support for stepping exercises and progressing to step touches with the intact limb. Upper limb support will progress from bilateral support to support provided only on the side of the prosthesis. Anterior and posterior stepping exercises will begin with the support of parallel bars while maintaining weight on the prosthesis.
Table 2. Level 2 device-specific physical therapy protocol.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Description</th>
<th>Criteria for advancement to level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>Reassessment as needed</td>
<td>N/A&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Strengthening</td>
<td>Progression to sitting, quadruped, planks, standing, and transitions</td>
<td>Independent in HEP&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Stretching</td>
<td>Primarily used for cooldown at completion of each session</td>
<td>Independent in HEP</td>
</tr>
<tr>
<td>Education</td>
<td>Description of HEP and purpose of each exercise</td>
<td>Ability to trigger power in prosthetic foot during gait &gt;50% of steps and increased gait symmetry with verbal cueing for step length as determined through observational gait analysis</td>
</tr>
<tr>
<td>Gait training</td>
<td>Improving step length symmetry (eg, verbal cueing to increase step length on nonprosthesis side) and improving rollover on prosthesis side (eg, resistive gait training with TheraBand)</td>
<td>Independent in HEP</td>
</tr>
<tr>
<td>Mobility training</td>
<td>Transfers (eg, stand-to-sit, and ramp negotiation)</td>
<td>Able to maintain midline center of mass with stand-to-sit transfer</td>
</tr>
<tr>
<td>Neuromuscular reeducation</td>
<td>Intact limb mobility (eg, rolling ball under intact limb) to promote weight shifting, static single limb balance training on prosthesis side with minimally necessary upper limb support and stepping exercises (eg, step touches 15-20 cm step in parallel bars and step touches to a cone in the parallel bars)</td>
<td>Independent in HEP</td>
</tr>
<tr>
<td>HEP</td>
<td>Includes therapeutic exercises outlined in level 2</td>
<td>Independent in HEP</td>
</tr>
</tbody>
</table>

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

<sup>b</sup>HEP: home exercise program.

**Level 3 (Sessions 5 and 6)**

At level 3 (Table 3), all strengthening will be performed exclusively in the HEP. PT will include neuromuscular reeducation progression, including multidirectional movements.

Stepping exercises will be performed with decreasing upper extremity support at a tolerance demonstrated by the participant maintaining an appropriate body position. Single limb stance activities will include perturbations, such as resistance with movements of a non–weight-bearing intact limb or standing on a foam pad or balance disc. Gait training will include multidirectional stepping with upper extremity support in the parallel bars. Resistive gait training will be introduced to promote proper mechanics for loading the prosthesis in stance and achieving maximal energy return at push off. In addition, manual proprioceptive neuromuscular facilitation will be performed to promote proper anterior pelvic rotation. Multidirectional ambulation will be progressed to outside of parallel bars.
**Table 3.** Level 3 device-specific physical therapy protocol.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Description</th>
<th>Criteria for advancement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 3 (sessions 5 and 6)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>Reassessment as needed</td>
<td>N/A&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Strengthening</td>
<td>Review as needed</td>
<td>Independent in HEP&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Stretching</td>
<td>To be used for cooldown at session completion</td>
<td>Independent in HEP</td>
</tr>
<tr>
<td>Education</td>
<td>Gait training and purpose for improving symmetry</td>
<td>Verbalizes understanding of gait training</td>
</tr>
<tr>
<td>Gait training</td>
<td>Promoting gait initiation on prosthesis side with anterior pelvic rotation</td>
<td>Demonstrates ability to trigger power with at least 80% accuracy during level ground ambulation and increased gait symmetry, including upright posture, step length, and toe break as determined through observational gait analysis</td>
</tr>
<tr>
<td>Mobility training</td>
<td>Ramp and stair negotiation</td>
<td>Hill Assessment Index score ≥6 (ie, step past more than half foot length, with assistive device)</td>
</tr>
<tr>
<td>Neuromuscular reeducation</td>
<td>Four-directional resistance exercise on intact side while maintaining single limb stance on prosthesis side, static single limb balance training on prosthesis side to be progressed to noncompliant surface (eg, foam), and progression of stepping exercises to increase time in single limb stand on the prosthesis</td>
<td>Able to maintain single limb stance on prosthesis side with or without an assistive device for 15 seconds</td>
</tr>
<tr>
<td>HEP</td>
<td>Includes therapeutic exercises in level 3</td>
<td>Independent in HEP</td>
</tr>
</tbody>
</table>

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

<sup>b</sup>HEP: home exercise program.

**Level 4 (Sessions 7 and 8)**

At level 4 (Table 4), PT will include advanced neuromuscular reeducation and gait training, followed by a final PT evaluation. Neuromuscular reeducation will include single limb squats in the parallel bars with upper extremity support, as needed. Gait training will continue resistive training on even surfaces, proprioceptive neuromuscular facilitation for pelvic rotation, verbal and tactile cueing for symmetrical and appropriate trunk rotation, and negotiation of uneven surfaces.

Individuals who do not meet the criteria to complete the PT program will be offered an additional 8 PT sessions after completion of the study.
Table 4. Level 4 device-specific physical therapy protocol.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Description</th>
<th>Criteria for advancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>Re-evaluation at final session</td>
<td>N/Aa</td>
</tr>
<tr>
<td>Strengthening</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Stretching</td>
<td>To be used for cooldown at session completion</td>
<td>Independent in HEPb</td>
</tr>
<tr>
<td>Education</td>
<td>Importance of continuation of HEP</td>
<td>Verbalizes understanding</td>
</tr>
<tr>
<td>Gait training</td>
<td>Promoting gait initiation on prosthesis side with anterior pelvic rotation, manual techniques and verbal cues to promote increased trunk rotation and trunk rotation symmetry, verbal cueing for symmetrical arm swing and trunk rotation (eg, “relax your shoulders”), and resistive gait training with TheraBand and verbal cues for increased step length with the intact limb and relaxed upright posture</td>
<td>Demonstrates ability to trigger power with at least 90% accuracy for powered prosthetic foot and increased gait symmetry for trunk rotation and arm swing, as determined through observational gait analysis</td>
</tr>
<tr>
<td>Mobility training</td>
<td>Ascending and descending ramps and stairs</td>
<td>Stair Assessment Index score of at least 4 (with assistive device, step-to pattern) for ascending and descending stairs, and ambulates in the community without increased participant-reported stumbles or falls compared with baseline</td>
</tr>
<tr>
<td>Neuromuscular reeducation</td>
<td>Applied during HEP</td>
<td>Demonstrates increased gait symmetry for step length, push off, anterior pelvic rotation, upright posture, trunk rotation, and arm swing</td>
</tr>
<tr>
<td>HEP</td>
<td>Review HEP</td>
<td>Demonstrates independence in HEP</td>
</tr>
</tbody>
</table>

aN/A: not applicable.
bHEP: home exercise program.

Data Collection Visit 2
After completion of the device-specific PT program, all participants will undergo data collection on the powered ankle-foot prosthesis. The participants will repeat the quality-of-life assessment, subjective surveys, pain assessment, cognitive load assessment, biomechanical gait analysis, and functional measures using the powered ankle-foot prosthesis, as described in the baseline visit. Following visit 2, participants will keep the powered ankle-foot prosthesis for an additional 4 weeks of home use and community use but will not undergo any further device-specific training.

Data Collection Visit 3
After the final 4 weeks of powered ankle-foot prosthesis use, participants will undergo final data collection. Participants will repeat the quality-of-life assessment, subjective surveys, pain assessment, cognitive load measurements, neurocognitive assessment, biomechanical gait analysis, and functional measures, as described in the baseline visit. Following data collection, all participants will be refitted with their energy storing and returning ankle-foot devices, and the powered ankle-foot devices will be returned to the study staff.

Statistical Analysis
Across the study population, outcomes will be assessed with descriptive statistics and compared between each ankle-foot device category as well as by PT intervention (ie, device-specific PT and standard of care). Inferential statistics for ordinal data will be conducted with a repeated-measures Friedman test (α=.05) and a Dunn post hoc test at a 95% CI. To address which measures are the most sensitive to intervention type, a linear mixed-effects model will be used. Separate models will be used for each type of measure (pain, subjective, cognitive, neurocognitive, functional, and biomechanical), and measures that have a significant association with the intervention type in the presence of adjusting (control) variables will be determined. Pair-wise comparisons will be tested for significance using linear contrasts with a Tukey honestly significant difference or by applying a Bonferroni correction, where applicable. The following sections outline the specific analyses that will be performed for each study objective.

Planned Statistical Analysis for Biomechanical and Functional Outcomes
Although there are numerous biomechanical and physiological parameters that can be evaluated following gait analysis [61], this investigation will focus on the biomechanical parameters that are most relevant, commonly used, able to discriminate, and have specific clinical relevance for individuals with transfemoral limb loss. The primary biomechanical outcome measures will include measures of rollover shape, individual characteristics of the 3D ground reaction force, and ankle, knee, and hip joint angles, moments, and powers (on both the intact and affected limbs). To evaluate the load distribution of the medial and lateral knee compartments, the peak resultant ground reaction force, ground reaction force rate, peak knee external adduction moments, and knee external adduction moment rate will be compared between the baseline (passive energy storing and returning condition) and the powered condition at each follow-up visit.
Linear mixed-effects models will be used to identify statistically significant differences in gait temporal-spatial and biomechanical variables for all walking speeds. The fixed effects will be the average differences in gait biomechanical and temporal-spatial variables by prosthetic ankle type (powered vs passive). These models also estimate random effects because of differences in mean biomechanical variables across participants, as well as the random effects associated with minimized variability, as the participants will be tested with both prostheses. For example, a linear mixed-effects regression will be used to examine the relationship between the intact knee peak external adduction moment and the prosthetic ankle-foot condition. Peak intact knee external adduction moment will be the dependent variable, whereas ankle-foot condition will be the independent variable, and participant-by-ankle-foot condition will be modeled as random effects. Pair-wise comparisons will be tested for significance using linear contrasts with a Tukey honestly significant difference or by applying a Bonferroni correction, where applicable. In addition, linear mixed-effects regression will be used to determine the association between prosthetic peak ankle power and foot condition.

**Planned Statistical Analysis for Pain Outcomes**

The following parameters will be measured and statistically compared between baseline (energy storing and returning condition) and each follow-up visit with the powered ankle-foot prosthesis (weeks 4 and 8):

- Joint reaction forces on the lower back (L5 and S1) and contralateral (intact) knees
- PEQ pain scores
- Functional outcome measures (6MWT, AMPpro, and CHAMP)

Spearman correlations will be calculated to correlate data between pain and lower extremity kinematic and kinetic parameters of interest and pain and functional outcome values across the groups. Linear mixed-effects models will be used to identify statistically significant differences in pain scores and biomechanical variables listed in the previous section for all walking speeds. For example, a linear mixed-effects regression will be used to examine the relationship between peak reaction moments at L5 and S1 and pain for each prosthetic ankle-foot condition. Pair-wise comparisons will be tested for significance using linear contrasts with a Tukey honestly significant difference or by applying a Bonferroni correction, where applicable.

Multiple linear regression will be performed with PEQ pain as the dependent variable. Because of the large number of predictor variables in relation to the number of participants, penalized methods (eg, ridge regression or lasso) will be used to identify variables that contribute the most to the prediction model. Penalized methods add a tuning parameter to the regression model that shrinks the less important coefficients toward 0. Cross-validation will be used to select the best tuning parameter value [62]. The independent variables will be comprised of functional parameters (eg, peak joint reaction forces) and the condition (ankle-foot type and intervention).

**Planned Statistical Analysis for Cognitive Load and Neurocognitive Outcomes**

Linear mixed-effects models will be performed to examine the differences between the ankle-foot devices on cognitive performance, walking speed, and subjective responses to attention. The parameters in the linear effects model will include prosthetic ankle-foot type, PT intervention, and participants. Ankle-foot type, PT intervention, and cognitive performance will be treated as fixed effects, whereas participants will be treated as random. The total number of errors and the error rates for the cognitive task will be calculated, and the mean error rate will be determined for each cognitive task performed by each participant. Repeated-measures ANOVA will be used to compare the error rates for the 3 cognitive tasks. Fisher least significant difference test will be used to make post hoc comparisons.

For the neurocognitive battery, Pearson correlation coefficients (pair-wise 2-tailed) will be calculated for all variables of interest. Stepwise multiple regression will be performed with the neurocognitive scores (Neurocognitive Index, composite memory, cognitive flexibility, and complex attention scores) as the dependent variable. The independent variables will comprise functional outcomes, pain, and gait biomechanics, including asymmetry index.

**Power Analysis and Sample Size Estimation**

The sample size was based on a power analysis of 3 biomechanical measures (leading limb work, ground reaction force rate, and knee adduction moment rate) and 1 functional outcome (6MWT distance), and 1 subjective outcome (PEQ—utility) obtained from preliminary analyses, with all measurements obtained at baseline and 2 additional measurements over an 8-week period. The sample size was calculated for the group-by-time interaction, which tests the differences in change over time between the study groups. Assuming an α error rate of 5%, Table 5 presents the power achieved for each measurement for 30 participants, 15 (50%) in each group.

---

https://www.researchprotocols.org/2024/1/e53412

JMIR Res Protoc 2024 | vol. 13 | e53412 | p.102

(page number not for citation purposes)
Results
This study was funded in September 2017, with enrollment beginning in September 2018. Data collection is expected to conclude by March 2024. Data analysis of the completed data set is expected to begin after final data collection. The initial dissemination of results is expected in August 2024, with subsequent publication of secondary analyses in December 2024.

Discussion
Expected Outcomes and Anticipated Principal Findings
After the completion of this research project, this investigation will have quantified the dependence of symmetrical lower limb gait biomechanics, physical function, and pain reduction on advanced prosthetic technology and device-specific rehabilitation. Furthermore, a device-specific treatment strategy designed to minimize impairments and maximize function will be evaluated. Finally, an objective measure of cognitive load and neurocognition to guide the prosthetic prescription of powered ankle-foot prostheses will be assessed. As such, the evidence-based outcomes obtained from this research investigation can be appropriately translated into clinical practice as well as drive the future of clinical care in this population.

VA provides care for veterans with limb loss of all generations, including the influx of service members with limb loss from the most recent conflicts [5]. Projected trends indicate that the overall number of individuals with limb loss will continue to increase dramatically, largely attributable to the aging population and the growing number of people with dysvascular disease and diabetes [1]. With this large population expected to grow, considerable resources will be required for rehabilitation and prosthetic services, driving limb loss care to become a high priority for VA. Effective outcomes-based clinical practice will be necessary to decrease long-term disabilities associated with prosthetic use and improve the quality of life. Therefore, it is the goal of this study to examine the effectiveness of a powered ankle-foot prosthesis and device-specific rehabilitation on gait biomechanics, performance, and pain in individuals with transfemoral limb loss. Results from this investigation will provide evidence-based outcomes that can be translated into successful strategies to minimize impairments and maximize function and may drive the evaluation of future advancements in prosthetic technology.

Dissemination Plan
The results of this investigation can help form evidence-based guidelines for individuals with transfemoral limb loss that can serve as a source for lower limb loss clinical practice guidelines. Dissemination of the results of this study within the DoD, VA, and the civilian health care systems will be performed in 3 ways. First, the results will be disseminated to the scientific and clinical community through traditional means, such as peer-reviewed submissions to professional conferences (eg, the Gait and Clinical Movement Analysis Society Annual Conference, the American Congress of Rehabilitation Medicine Annual Conference, and the Military Health System Research Symposium), targeted limb loss and rehabilitation publications (eg, Archives of Physical Medicine and Rehabilitation, Gait & Posture, and Frontiers in Bioengineering and Biotechnology), as well as sharing the data through large data repositories. Second, both VA and the DoD have national, interdisciplinary groups and committees that enable the national dissemination and adoption of best practices among different disciplines. For the VA and DoD limb loss care teams, the Extremity Trauma and Amputation Center of Excellence and the VA Amputation System of Care hold a bimonthly webinar series for clinicians, scientists, and researchers that is available across the entire DoD and VA health care network. This series allows research results to be presented to a large, diverse audience of researchers and health care professionals in limb loss care, which can directly influence the care provided to veterans and service members with limb loss. Finally, the outcomes will be disseminated directly to leaders in the prosthetics industry to provide real-world feedback on their products. The results provided to industry leaders can help in the evolution of lower extremity prosthetic components, which can then lead to improved devices for individuals with transfemoral limb loss.

Limitations
This investigation will not address the varied dosing, timing, frequency, and duration of the device-specific rehabilitation protocol. However, the frequency of PT visits and protocol timing will be evaluated for each participant, which can provide preliminary data for a future study to optimize the rehabilitation strategy. The heterogeneity of the sample population may also limit the generalizability of the outcomes to a more diverse population. In addition, the statistical analysis models will also adjust for specific parameters (eg, age, time since limb loss, and etiology of limb loss), which may limit the sample size and interpretability of the results. The type of microprocessor knee

<table>
<thead>
<tr>
<th>Measure</th>
<th>Difference, mean (SD)</th>
<th>Power (N=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leading limb work  (J/kg)</td>
<td>0.03 (0.04)</td>
<td>85</td>
</tr>
<tr>
<td>Ground reaction force rate (N/kg/s)</td>
<td>15 (26.9)</td>
<td>82</td>
</tr>
<tr>
<td>Knee adduction moment rate (Nm/kg/s)</td>
<td>1.1 (1.2)</td>
<td>98</td>
</tr>
<tr>
<td>6MWT  distance (m)</td>
<td>60 (61)</td>
<td>96</td>
</tr>
<tr>
<td>PEQ—utility</td>
<td>8.4 (8.8)</td>
<td>96</td>
</tr>
</tbody>
</table>

*a*6MWT: 6-minute walk test.

`PEQ: Prosthesis Evaluation Questionnaire.`
used by each participant will not be prescribed and may be a confounding factor. This influence of the microprocessor knee type will be evaluated in the statistical models. Finally, a 4-week assessment following completion of the device-specific protocol may not be sufficient to evaluate any PT rebound effects or long-term changes in gait, specifically regarding the effectiveness of the acute PT intervention on gait biomechanics.

Acknowledgments
At Veterans Affairs New York Harbor Healthcare System, the authors would like to acknowledge Mr Michael Poppo for his assistance during data collection and processing; Mr Eric Baksh; Mr Kenneth Breuer; and Mr Eric Hubbs for their assistance in prosthesis fitting and alignment and Dr Cristina Parker, Dr Ellen Godwin, and Dr Jonathan Glasberg for conducting the rehabilitation program. At Walter Reed National Military Medical Center, the authors would like to acknowledge Mr Jonathan Gladish for his assistance during data collection and processing, Dr Ashley Knight for her assistance in project administration, and Dr Allison Symssack for helping develop the device-specific physical therapy program. At James A. Haley Veterans' Hospital, the authors would like to acknowledge Dr Meghan Kern and Dr Lisa Ballestrea for their assistance during data collection and participant training and Ms Anh Du for her assistance in prosthesis fitting and alignment. Finally, the authors would like to acknowledge the Narrows Institute for Biomedical Research and Education for their assistance in grant management. The Congressionally Directed Medical Research Programs, Orthotics, and Prosthetics Outcomes Research Program provided support (grant W81XWH-17-1-0568). This funding source 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.

Disclaimer
The views expressed in this paper are those of the authors and do not necessarily reflect the official policy of the departments of the Air Force, Army, Navy, Defense, Veterans Affairs, or the United States government. The identification of specific products or scientific instrumentation is considered an integral part of the scientific endeavor and does not constitute endorsement or implied endorsement on the part of the author, the Department of Veterans Affairs, the Department of Defense, or any component agency.

Data Availability
All data for this manuscript are included in this published paper. Final results will be available via a publicly available data repository.

Authors' Contributions
JTM, ALP, and LMN participated in conceptualization of the study. JTM, ALP, BDH, and CLD participated in methodology development. JTM, ALP, BDH, DVH, JMC, MIH, SLP, and ANS participated in data collection. JTM, ALP, BDH, DVH, JMC, MIH, SLP, ANS, CLD, and LMN participated in writing, editing, and review of the manuscript. JTM, BDH, MIH, and SLP participated in project administration. JTM, BDH, and SLP participated in supervision of the project and staff. JTM, ALP, CLD, and LMN participated in funding acquisition. All authors read and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Peer-review summary statement.
[PDF File (Adobe PDF File), 121 KB - resprot_v13i1e53412_app1.pdf ]

Multimedia Appendix 2
Peer-review response from the investigators.
[PDF File (Adobe PDF File), 126 KB - resprot_v13i1e53412_app2.pdf ]

References


Abbreviations

6MWT: 6-minute walk test
AMPpro: Amputee Mobility Predictor with prosthesis
CHAMP: Comprehensive High-Level Activity Mobility Predictor
DoD: Department of Defense
HEP: home exercise program
IRB: institutional review board
JAHVH: James A. Haley Veterans’ Hospital
PEQ: Prosthesis Evaluation Questionnaire
PT: physical therapy
VA: Veterans Affairs
VANYHHS: Veterans Affairs New York Harbor Healthcare System
WRNMMC: Walter Reed National Military Medical Center
Exploring Medical Career Choice to Better Inform Swiss Physician Workforce Planning: Protocol for a National Cohort Study

Milena Abbiati¹, PhD; Mathieu R Nendaz¹,², MD, MHPE; Bernard Cerutti¹, MPH, PhD; Monika Brodmann Mäder³,⁴, MD, MME; Giatgen A Spinas³, MD; David Vicente Alvarez⁵, MS; Douglas Teodoro⁵, PhD; Georges L Savoldelli¹,⁶, MD, MEd; Nadia M Bajwa¹,⁷, MD, MHPE, PhD

¹Unit of Development and Research in Medical Education, Faculty of Medicine, University of Geneva, Geneva, Switzerland
²Division of Internal Medicine, Department of Medicine, Geneva University Hospitals, Geneva, Switzerland
³Swiss Institute for Medical Education, Bern, Switzerland
⁴Department of Emergency Medicine, University Hospital Bern, Bern University, Bern, Switzerland
⁵Department of Radiology and Medical Informatics, Faculty of Medicine, University of Geneva, Geneva, Switzerland
⁶Division of Anesthesia, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland
⁷Division of General Pediatrics, Department of Women, Children, and Adolescents, Geneva University Hospitals, Geneva, Switzerland

Corresponding Author:
Milena Abbiati, PhD
Unit of Development and Research in Medical Education
Faculty of Medicine
University of Geneva
Centre Médical Universitaire
1 Rue Michel-Servet
Geneva, 1206
Switzerland
Phone: 41 22 379 59 63
Fax: 41 22 379 51 22
Email: Milena.Abbiati@unige.ch

Abstract

Background: A medical student’s career choice directly influences the physician workforce shortage and the misdistribution of resources. First, individual and contextual factors related to career choice have been evaluated separately, but their interaction over time is unclear. Second, actual career choice, reasons for this choice, and the influence of national political strategies are currently unknown in Switzerland.

Objective: The overall objective of this study is to better understand the process of Swiss medical students’ career choice and to predict this choice. Our specific aims will be to examine the predominately static (ie, sociodemographic and personality traits) and predominately dynamic (ie, learning context perceptions, anxiety state, motivation, and motives for career choice) variables that predict the career choice of Swiss medical school students, as well as their interaction, and to examine the evolution of Swiss medical students’ career choice and their ultimate career path, including an international comparison with French medical students.

Methods: The Swiss Medical Career Choice study is a national, multi-institution, and longitudinal study in which all medical students at all medical schools in Switzerland are eligible to participate. Data will be collected over 4 years for 4 cohorts of medical students using questionnaires in years 4 and 6. We will perform a follow-up during postgraduate training year 2 for medical graduates between 2018 and 2022. We will compare the different Swiss medical schools and a French medical school (the University of Strasbourg Faculty of Medicine). We will also examine the effect of new medical master’s programs in terms of career choice and location of practice. For aim 2, in collaboration with the Swiss Institute for Medical Education, we will implement a national career choice tracking system and identify the final career choice of 2 cohorts of medical students who graduated from 4 Swiss medical schools from 2010 to 2012. We will also develop a model to predict their final career choice. Data analysis will be conducted using inferential statistics, and machine learning approaches will be used to refine the predictive model.

Results: This study was funded by the Swiss National Science Foundation in January 2023. Recruitment began in May 2023. Data analysis will begin after the completion of the first cohort data collection.
Conclusions: Our research will inform national stakeholders and medical schools on the prediction of students’ future career choice and on key aspects of physician workforce planning. We will identify targeted actions that may be implemented during medical school and may ultimately influence career choice and encourage the correct number of physicians in the right specialties to fulfill the needs of currently underserved regions.

International Registered Report Identifier (IRRID): DERR1-10.2196/53138

(JMIR Res Protoc 2024;13:e53138) doi:10.2196/53138

KEYWORDS

career choice; medical specialty; medically underserved area; motivation; professional practice; medical students; residents; prediction model; machine learning; physician workforce

Introduction

Background

The unbalanced distribution of medical staff is identified by the Global Health Workforce Alliance as the major public health challenge of the 21st century [1]. The equitable distribution of physicians in the appropriate specialties and in the appropriate regions challenges high-income countries that are currently dependent on foreign-trained physicians. The increasing pressure to supplement the physician workforce is because of the growing medical needs of an aging population coupled with an aging physician workforce [2,3]. A misdistribution of the physician workforce has been clearly predicted in recent statistics owing to increasing rates of part-time employment, difficulties in recruiting physicians to certain specialties, and growing gender gaps in certain specialties and places of practice [4,5].

This is particularly true in Switzerland, where life expectancy is high and 1 of 4 physicians is currently aged >60 years [2,3]. To increase the number of physicians, the federal and cantonal states have financed the creation of 4 new medical schools. However, there is little link, if any, between the physician and general population ratio and the distribution of the medical workforce; therefore, this strategy may be largely insufficient to meet the needs of the population [6,7].

Medical schools and postgraduate medical education policymaking bodies play a key role in this dynamic. They are responsible for the supply of an adequate medical workforce to meet demands for quality, quantity, and appropriate distribution of physicians among specialties and geographic areas [8-10]. The regulation of the Swiss medical postgraduate training system offers great flexibility to physician graduates regarding specialty selection, postgraduate program duration, and location. It is unique and differs from other systems in neighboring countries, where the specialty choice and training site are based on numerus clausus [2-4]. Beyond the regulation system, there is often an imbalance between students’ wishes, the need for specialists, and the positions available, leading to shortages in some specialties and high levels of competition in others [11]. Furthermore, once specialization is achieved, concentrations in certain types of practice (private outpatient rather than hospital based) and in certain geographic areas (advantaged urban rather than rural or disadvantaged urban areas) are unbalanced [12]. Consequently, there is a growing effort to better understand what factors contribute to the physician workforce shortage and misdistribution.

These factors are impacted by the societal, political, and regulatory aspects of postgraduate medical training. The Bland-Meurer theoretical framework for medical career decision-making indicates that both specialty characteristics (eg, type of practice and person oriented) and students’ career needs (eg, prestige and work-life balance) are shaped by students’ individual static (eg, demographic) and dynamic (eg, motivation) features, by life experiences, and by the values and culture of the training institution [13,14]. Studies on predominately dynamic characteristics, such as motivational factors, show that students’ specialty choice can be driven by internal motives ("intrinsic motivation"), such as personal abilities, interest in helping patients, intellectual challenge, and by external motives ("extrinsic motivation"), such as salary, status, and workload [15]. Their importance differs by specialty [16,17] and gender [18-21].

Various methods of artificial intelligence and machine learning have also been investigated to provide decision support for and predict individual career choice [22-28]. These career choice prediction and decision support studies have achieved some success at being able to predict career choice, with performance varying from 68% to 85% area under the receiver operating characteristic curve. To the best of our knowledge, only 1 study has used modern machine learning methods to predict a medical student’s career choice [29]. To date, it remains unknown how these models perform in career choice analyses in which individual and contextual data are available. Moreover, how machine learning models perform in predictive scenarios for multiple medical career choices is yet to be explored.

Current State of Our Own Research

In our previous research, we explored the areas of intention to practice (urban and rural) of graduating students at 4 Swiss medical schools [30]. We found that 13.7% expressed an intention to practice in underserved areas (62.1% of whom intended to practice in rural areas) and 41.1% were undecided. These intentions varied from one school to another and were related to different motivational factors. Work variety and work conditions appear to be factors that might attract interested and undecided students to work in underserved areas. Among those who wished to practice in underserved areas, general practice (21.6%) was the most preferred specialty. Motivational factors influencing specialty choice were intellectual challenges, work variety, work conditions, and enthusiasm. In addition, using the same cohort at 4 Swiss medical schools, we analyzed the motivating factors that influence the choice of obstetrics and gynecology as career intentions. The results highlighted the
importance of “experiential factors” and gender in this specialty choice. These findings provide useful information for targeted interventions to promote obstetrics and gynecology during undergraduate and postgraduate training by providing more hands-on experiences and improving the integration of male students and residents [31]. We also explored the degree of motivation for general practice, surgery, radiology, and psychiatry throughout the preclinical years. We specifically focused on personal and motivational individual characteristics correlated with general practice and surgical career intention [32-35].

The Swiss Medical Career Choice study is the continuation of 2 preliminary studies [12,36]. We will explore the role of students’ individual static and dynamic features as well as that of the educational organizational context and training institution culture on career choice, that is, specialty, type, and place of practice, from medical school to postgraduate training. Thus, the present proposal seeks to understand the process of Swiss medical student career choice and attempts to predict this choice. First, we will examine the static and dynamic variables that predict the career choice of Swiss medical school graduates and their interaction. Second, we will examine the evolution of Swiss medical students’ career choice and their ultimate career path.

Study Objectives

Primary Objectives

The primary objectives are to determine the current career intentions of Swiss medical students (choice of specialty and type of practice) and to assess the personal and contextual factors that determine their choice.

Secondary Objectives

The secondary objectives are as follows:

- To assess the influence of static factors such as gender and socioeconomic status
- To assess the influence of dynamic factors such as student perceptions of medical specialties and motivation
- To assess the interactions between both types of factors
- To attempt to predict career choice based on static, dynamic, and motivational factors and to design and test data-driven methods (artificial intelligence) to predict medical students’ career choice
- To determine how career choice intentions vary during medical school within and across different medical school sites and during postgraduate training
- To determine how career choice intentions of Swiss medical school graduates have evolved over the last decade, considering the political strategies that have been put into place
- To compare how career choice intentions differ from the final choice in the Swiss nonregulated system and the French regulated system

Methods

Study Design

The Swiss Medical Career Choice is a 24-month longitudinal prospective national investigation implemented over 4 years (2 data collection time points) for 4 cohorts of medical students across all Swiss medical schools as well as a follow-up during postgraduate training.

For comparison purposes, this study also includes (1) a follow-up of 2 previous cohorts of medical students from Western Switzerland during postgraduate training, (2) a follow-up of the final career choice of 2 previous cohorts of medical students from 4 Swiss medical schools, and (3) a follow-up of the final career choice of 2 cohorts of medical students from Strasbourg, France.

Participants

In Switzerland, medical school consists of a 6-year curriculum divided into 3 years of bachelor’s (basic science training) and 3 years of master’s (clinical training). For the national prospective study, eligible participants will be medical students entering year 4 (master’s 1) between 2022 and 2024 and finishing year 6 (graduates) between 2023 and 2026. The total number of eligible participants is approximately 1440 per cohort. Participants who have completed the questionnaire either in year 4 or 6 and have a response rate of >90% on the structured scale will be eligible for the follow-up study in postgraduate training year 2.

Data Collection

Data will be collected through 4 different sources: (1) a questionnaire administered during undergraduate year 4 (master’s 1) and year 6 (master’s 2), (2) a short survey during postgraduate training year 2, (3) data extracted from the residents’ logbook of the Swiss Institute for Medical Education (SIME), and (4) exogenous contextual data extracted from the Federal Offices of Public Health and Statistics and the Swiss Medical Association (FMH).

Undergraduate participants will be invited to complete the questionnaire (approximately 30 min) during a compulsory class at the beginning of year 4 and at the end of year 6. Data will be anonymized for confidentiality reasons and for data protection issues. Participants will provide their student ID for matching purposes for the duration of the study. During the postgraduate years, participants (residents) will be invited to complete the short questionnaire (approximately 10 min) by invitation from the SIME, and relevant data will be extracted from their logbook. Data matching between the 2 data sources (undergraduate survey and data collected during residency) will be performed for each participant based on their ID number.

Ethical Considerations

The Chair of the Cantonal Commission for Ethical Research designated this study as exempt from formal review (protocol BASEC 2020-00813) as the aim of the study aim is outside the scope of the Swiss law as defined in Article 2 of the Human Research Act (HRA).
To obtain informed consent for undergraduate data collection, eligible participants will receive an email 10 days before the survey to inform them about the research project’s main goals, the content, and the testing conditions (confidential and voluntary participation). Students who agree to participate will confirm their informed consent by marking a box on the first page of the questionnaire. Comparison cohorts of medical students consented previously to this study.

Data matching will be performed by a technical administrator who is not involved in the data analysis and interpretation. Researchers will only have access to deidentified data. Anonymous responses will be collected and stored on the secure web-based server Evasys. Survey data will be extracted from Evasys and stored in a password-protected Excel (version 21; Microsoft Corporation) file that is accessible only to the central study team. The University of Geneva standards for data handling will be followed for all data management and record keeping.

Participants will not receive compensation for completing the questionnaire.

Measures

Main Outcomes

The intention of practice of the undergraduates will be assessed through a single-choice question among 6 possible options grouped into four categories: (1) hospital-based medicine (senior physician in a nonuniversity public hospital and academic and clinical career in a university hospital), (2) office-based medicine (private clinical practice in a solo practice and private clinical practice in a group practice), (3) research and teaching, and (4) undecided. The students’ specialty intentions will be gathered through a single-choice question among the 46 federal specialist titles issued by the SIME plus geriatrics, emergency medicine, and an undecided option. Specialties will be further regrouped into seven categories of intentions: (1) surgical, (2) acute care, (3) diagnostic medicine, (4) preventive medicine, (5) medical subspecialties, (6) general practice, and (7) undecided. Categories 1, 2, and 3 will be grouped together in the supracategory of “technically oriented specialties.” Categories 4, 5, and 6 will be grouped together in the supracategory “person-centered specialties.” Data collected during the postgraduate training and extracted from the SIME residents’ logbook include the specialty in which the residents are currently registered, specialist title for which the resident is aiming, number of months completed in the desired specialty, and the number of months completed in different specialties. The index of change during medical school or residency will be calculated as follows: if the student does not change=0, if the student changes their intention in the same specialty category=0.5, and if the student changes for another supracategory=1. The frequency of changes will also be calculated.

Demographic Data

The demographic will include age, gender, nationality (Swiss, European, or other), medical school, high school diploma (scientific vs other), place of origin, marital status, mother tongue, and 2 indirect measures of students’ socioeconomic level, that is, parents’ highest educational achievement (primary, secondary, or tertiary), and parents’ profession (elementary, employee, executive, or professional) [37,38].

Motivational Factors

A total of 5 global motivational factors identified in the scientific literature as influencing the choice of specialty, validated by Beaulieu et al [36], will be ranked by their importance. The degree of motivation to become a surgeon as well as a general practitioner will be measured on a 6-point Likert scale: 1=very unmotivated to 6=very motivated. The global motivational factors mentioned earlier will be broken down into 12 specific motivational factors [34] influencing the choice to become a surgeon and a general practitioner (measured on a 6-point Likert scale: 1=very dissuasive to 6=very attractive). In total, 2 single-choice questions will assess (1) students’ intention to practice in medically underserved areas (yes, undecided, or no) and (2) if yes, they will be asked to specify the desired location of practice (rural, mountain, or urban areas). The desired percentage of employment will be measured as the percentage of full-time employment on a 10-point Likert scale (0=0% to 10=100%).

Individual Characteristics

The previous academic background will be identified through 2 items asking students to report the type of high school degree and the grade obtained. Personality traits will be measured through the NEO Five Factor Inventory, which is widely used to assess the major personality traits as described in the Big Five Model [39]. It consists of 60 items, 12 per trait, scored on 5-point Likert scales (0=strongly disagree to 4=strongly agree). Motivation will be measured through the Academic Motivation Scale, which is widely used to assess self-determinate motivation in educational settings [40]. It consists of 28 items scored on a 7-point Likert scale (1=strongly disagree to 7=strongly agree) assessing intrinsic motivation (12 items), extrinsic motivation (12 items), and amotivation (4 items). Students’ anxiety will be measured through the State-Trait Anxiety Inventory Form Y-State. Notably, we will assess state anxiety using the state anxiety subscale [41]. It consists of 20 items scored on a 4-point Likert scale (1=strongly disagree to 4=strongly agree). The total score ranges from 20 to 80. Scores are categorized into a 3-point cutoff: below 55 (average anxiety), 56 to 65 (high anxiety), and above 65 (severe anxiety). Students’ gender representation will be measured through 6 items assessing gender bias [42]. Single-item scores will be standardized into t scores, and a total score will be calculated.

Context Characteristics

Students’ perception of learning context will be measured using a brief revised version of the students’ perception of teachers in the Dundee Ready Educational Environment Measure 11-item subscale [43]. This revised version consists of 6 items scored from 0 (strongly disagree) to 4 (strongly agree) as in the original version (max score of 24). The selected questions will assess if students have identified a person they view as a role model (yes, during my medical school training; yes, but not during my medical school training; and no). If yes, 2 open-ended questions will assess (1) the function of this person and (2) in what context they met them. A total of 2 single-choice questions will assess

https://www.researchprotocols.org/2024/1/e53138

JMIR Res Protoc 2024 | vol. 13 | e53138 | p.112

(page number not for citation purposes)
(1) if the actual program location is the desired one and (2), if not, which location is the desired one. In total, 2 6-point Likert scale questions will assess (3) the degree of importance of the language of instruction and (4) the degree of importance of the cultural context.

Exogenous data collected from the Federal Medical and Population Statistics and the FMH include population density and medical density (of area of origin and of medical schools), number of professionals per specialty, specialist density per region, average specialist salary, and average number of working hours.

Table 1 summarizes domains, constructs, and measures by time points of data collection.

Table 1. Overview of domains, constructs, and measures by collection time points.

<table>
<thead>
<tr>
<th>Domain and construct</th>
<th>Measure</th>
<th>Collection time point(^a)</th>
<th>(^b)</th>
<th>(^c)</th>
<th>(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type or category of practice</td>
<td>Single-choice question</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Type or category of specialty</td>
<td>Single-choice question</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rate of specialty or practice change</td>
<td>Homemade scale</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Age, gender, and other demographics</td>
<td>Single-choice questions</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Marital status</td>
<td>Single-choice question</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Motivational factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global motives for career choice</td>
<td>6-point Likert scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Motivation for general practice and surgery</td>
<td>6-point Likert scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Motives to choose general practice or surgery</td>
<td>6-point Likert scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Intention to practice in underserved areas</td>
<td>Single-choice question</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Desired percentage of employment</td>
<td>9-point Likert scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Individual factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous academic background</td>
<td>Single-choice question</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Personality</td>
<td>NEO-FFI(^f) [39]</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Motivation</td>
<td>AMS(^g) [40]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anxiety</td>
<td>STAI-A(^h) [41]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gender bias</td>
<td>Homemade scale</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Context</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning context</td>
<td>DREEM-R(^i) [43]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Role modeling</td>
<td>Single-choice questions</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Master’s or training program location</td>
<td>Single-choice questions</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Exogenous data</td>
<td>Official records</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

\(^a\) The questionnaires will be collected annually from 2023 to 2026 from all fourth-year and sixth-year medical students and residents of Switzerland.

\(^b\) Master 1: fourth year of medical school.

\(^c\) Master 3: sixth year of medical school.

\(^d\) PGY: postgraduate year.

\(^e\) Data collection for the associated variable at the time point.

\(^f\) NEO-FFI: NEO Five Factor Inventory.

\(^g\) AMS: Academic Motivation Scale.

\(^h\) STAI-A: State-Trait Anxiety Inventory Form A.

\(^i\) DREEM-R: Dundee Ready Educational Environment Measure.

In the postgraduate survey, the selected items of individual characteristic measures will be used, except for the State-Trait Anxiety Inventory Form A. Data collection for the Western Switzerland and French comparison cohorts is similar to that...
for the prospective national cohort. For the Western Switzerland cohorts, we will conduct 2 cross-sectional data collections at 2 different time points. For all comparison cohorts, we will retrieve the career choice through the residents’ logbook of the SIME and through the National Examination (Epreuves Nationales) and the National Classifying Examinations (Epreuves Classantes Nationales) official site in France.

Data Analysis

This study adheres to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cohort studies [44].

Descriptive statistics will be applied to demographics, career intention, motivational factors, and individual and contextual characteristics. A chi-square test will be used to compare intentions by medical school as well as to compare previous cohorts with the national prospective cohorts. The type I error rate will be set at .05.

To examine specific motivational factors, a principal component analysis with varimax rotation will be run to aggregate the motives out of the 12, demonstrating a frequency of >10%. Data suitability will be confirmed using the Kaiser-Meyer-Olkin index of sampling adequacy. Combined criteria (ie, scree plot, eigenvalue >1.0, and interpretability) will be used to determine the number of factors [45]. The critical value for significant factor loading will be >0.40 [46]. Each factor obtained from the principal component analysis procedure will be labeled according to its content.

Appropriate statistical tests will be applied to measure changes in motivational factors, the Academic Motivation Scale, anxiety, and contextual characteristics between different time points, such as rank-order stability (Spearman correlation ρ coefficient), the Cohen d effect size magnitude ratios [47], and the Reliable Change Index [48].

To examine whether the demographic, motivational factors, and individual and contextual characteristics will correlate with our main outcomes, logistic regression (odds ratio and 95% CIs) and linear regression will be used. These statistical analyses will be performed using R (version 4 or above; R Foundation for Statistical Computing) and SPSS (version 27; IBM Corp).

Following a machine learning approach, the questionnaire data and exogenous contextual data from the Federal Medical and Population Statistics and the FMH will be combined in an attempt to obtain a comprehensive view of the different medical careers within the working environment. This combined data set will then be preprocessed according to the input format of the machine learning models and will link medical students and exogenous contextual data to their career intentions. To avoid issues of data bias being injected into the algorithms, data augmentation via downsampling and oversampling strategies will be assessed to reflect equal gender distribution in the training set. For the model design and experiments, different machine learning algorithms will be investigated, in particular extreme gradient boosting [49] and CatBoost [50], which provide state-of-the-art performance for categorical data, to learn career choice patterns from the collected data set. In our experiments, the data set will be randomly divided into training (60%), development (20%), and test (20%) sets to train the model parameters and hyperparameters and to evaluate the models’ performance, respectively. Hyperparameter tuning and evaluation metrics will be computed using cross-fold validation to increase the robustness of the results. Standard classification evaluation metrics such as area under the receiver operating characteristic curve, $F_1$-score, precision, and recall will be reported. The comparison of models’ predictive results will be measured using the McNemar statistical test. The type I error rate will be set at .05. The Shapley additive explanation method [51], a consistent, fast, and deterministic method for extracting feature contributions at the individual prediction level, will be used to identify factors impacting career choices.

Statistical Power

Regarding the planned analyses, all the main outcomes deal with the estimation of proportions (eg, specialty choice and practice in underserved areas). Considering the smallest subgroup population of the study (2 cohorts of residents, ie, approximately n=2880) and the proportion associated with the highest variability of the estimates ($P=.50$), a sample size of 864 (ie, 2×432; refer to Table 1) would allow to estimate any proportion with a precision of +2.790% to −2.790% (95% CIs derived from the hypergeometric distribution).

Results

The project has been peer reviewed and funded by the Swiss National Science Foundation in November 2022 with a start date of January 1, 2023, and an end date of December 31, 2026 (Multimedia Appendix 1). Data collection is currently underway, with the longitudinal cohort study having launched nationally on March 31, 2023. Because of organizational constraints, the postgraduate follow-up of historical cohorts will be launched from autumn 2023 to winter 2024. We expect to obtain preliminary results by mid-2024.

Discussion

Overview

This paper describes a longitudinal, prospective national investigation that will survey 4 cohorts of medical students across all Swiss medical schools as well as residents during their initial postgraduate training. The main objectives are to better understand the career choice process of Swiss medical students and to try to predict this choice.

This project will allow us to better understand the individual and contextual factors influencing Swiss medical students’ career path, from students’ intention at the end of medical school to postgraduate medical training and final specialization. This study will deepen preliminary findings on the relative influence and interaction of static and dynamic variables such as gender, work-life balance, and students’ medical specialty perceptions [12,32-36]. The fact that the study is conducted in all Swiss medical schools will also pinpoint the differences between the various Swiss medical schools and, in particular, the effects of new medical master’s programs regarding career choice and location of practice. This study will allow us to inform national stakeholders and medical schools both through prediction of

https://www.researchprotocols.org/2024/1/e53138
students’ future choices and key aspects of physician workforce planning. Finally, by drawing comparisons with a regulated postgraduate training system such as France’s, this study will also have potential benefits at the international level.

Strengths
To the best of our knowledge, this study is the first to investigate the personal characteristics and the evolution of students’ career choices during their medical studies and postgraduate training at a national level. Similarly, there is a lack of data regarding the follow-up from the initial intention to the definitive choice. This information is essential because better monitoring and understanding of career choice paths could help promote the management of physician resources and direct undergraduate and postgraduate interventions aimed at a better distribution of these resources [8-10].

We will provide information and prediction tools to meet demands for quality, quantity, and appropriate distribution of physicians among specialties and among geographic areas.

Using both individual and contextual data from medical students, we should be able to improve the predictive performance of the machine learning models as compared with questionnaire-only data [29]. By enabling better estimates than standard past averages, we expect that the predictive models will provide more effective support to decision makers for capacity planning. Moreover, by identifying the factors impacting career choice, decision makers will have data-driven information to support mitigation actions against specialty shortage. Finally, the predictive models may also help medical students in their career choice.

Limitations
Difficulties in recruiting collaborators within each medical school and coordination for survey administration may hinder response rates from students and therefore impact the realization of the study’s objectives. Finally, most of the variables are self-reported and therefore subject to personal bias. However, this study will use questionnaires with established evidence to support their validity. In addition, the project has received official support from the Joint Commission of Swiss Medical Schools. This allows us to include students from all medical schools and from multiple years of study. The partnership with the Swiss postgraduate governing body, SIME, is key to enabling data collection and follow-up of subjects during the postgraduate training years. The interdisciplinary research team brings together a psychologist, a statistician, a computer engineer, and 5 medical doctors, all involved in medical education at the undergraduate and postgraduate level and holding different specialty titles.

Conclusions
Exploring the individual and contextual factors associated with the career path of all medical students in Switzerland and the establishment of a follow-up system will provide important information for improving the quality of medical workforce planning.

We will identify targeted actions that may be implemented during medical school and may ultimately influence career choice and encourage the correct number of physicians in the right specialties to fulfill the needs of currently underserved regions. Potentially, these results could contribute to better management of the medical workforce by balancing future physician distribution and, in turn, increasing the efficiency of the health care system and meeting the needs of Swiss society.

Acknowledgments
This work is supported by the Swiss National Science Foundation (grant 100019_213172/1). The Swiss National Science Foundation played no role in the creation, review, or approval of this manuscript for publication. The authors would like to thank Zoya Horcik, Dr Philippe Conus, Dr David Gachoud, Dr Régis Marion-Veyron, Dr Julie Palix, and Dr Romano Regazzi of the University of Lausanne, Dr Michele Ghielmini, and Dr Marilù Guigli-Poretti of the University of Italian-speaking Switzerland, Dr Jean Sibilia and Dr Thierry Pelaccia of the University of Strasbourg, and Dr Raphael Bonvin and Dr Elke Bayha of the University of Fribourg for their support in the initial implementation of the study at their respective universities. The authors would like to thank Eva Hennel and Lukas Wyss for their help with the postgraduate aspect of their current collaboration with the Swiss Institute for Medical Education.

Data Availability
The data sets generated and/or analyzed in preparation for / prior to this study that underpin a publication are available from the corresponding author on reasonable request. The data sets generated and/or analyzed as part of this study that underpin a publication will be available in due course at the Yareta, University of Geneva data repository.

Authors’ Contributions
MA, MRN, DVA, DT, GLS, and NMB contributed to the initial draft of the manuscript, whereas BC, MBM, and GAS reviewed, revised, and approved the final submission. GLS and NMB contributed equally to this manuscript.

Conflicts of Interest
None declared.
References


Abbreviations

FMH: Swiss Medical Association
SIME: Swiss Institute for Medical Education
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

©Milena Abbiati, Mathieu R Nendaz, Bernard Cerutti, Monika Brodmann Mäder, Giatgen A Spinas, David Vicente Alvarez, Douglas Teodoro, Georges L Savoldelli, Nadia M Bajwa. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.01.2024. 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

Participatory Development and Assessment of Audio-Delivered Interventions and Written Material and Their Impact on the Perception, Knowledge, and Attitudes Toward Leprosy in Nigeria: Protocol for a Cluster Randomized Controlled Trial

Ngozi Murphy-Okpala¹, MBBS, FMCPH; Tahir Dahiru², MBBS, MPH; Anna T van ’t Noordende³, MSc, PhD; Carolin Gunesch⁴, MA; Joseph Chukwu¹, MBBS, MPH; Charles Nwafor¹, MBBS, MPH; Suleiman Hudu Abdullahi², MBBS; Chukwuma Anyaike⁵, MBBS, DTM&H, MCommH, FWACP; Ugochinyere Angelic Okereke⁶, NCE, BAEd, ME; Anthony Meka¹, MBBS, MPH; Chinwe Eze⁴, BSc, MPH; Okechukwu Ezeakile¹, MBBS, MPH; Ngozi Ekeke¹, MBBS, MPH, FWACP

¹RedAid Nigeria, Enugu, Nigeria
²Leprosy and Tuberculosis Relief Initiative, Jos, Nigeria
³NLR International, The Netherlands, Amsterdam, Netherlands
⁴DAHW-German Leprosy and TB Relief Association, Würzburg, Germany
⁵National Tuberculosis and Leprosy Control Programme, Federal Ministry of Health, Abuja, Nigeria
⁶Department of Art Education, University of Nigeria, Nsukka, Nigeria

Corresponding Author:
Ngozi Murphy-Okpala, MBBS, FMCPH
RedAid Nigeria
56 Nza Street, Independence Layout
Enugu, 23401
Nigeria
Phone: 234 706 749 0052
Email: ngozi.murphyokpala@redaid-nigeria.org

Abstract

Background: In Nigeria, similar to many leprosy-endemic countries, leprosy is highly stigmatized. High levels of stigma among community members as well as internalized stigma among persons affected by leprosy often result in negative psychosocial consequences for those affected. To break this vicious cycle, it is important to conduct context-specific behavioral change activities. Although written material has been successful in improving knowledge and perception, it is not suitable for populations with low educational levels. Audio-delivered interventions are likely to be more suitable for people who are illiterate. This study proposes to assess the impact of an audio-delivered intervention on the perception (knowledge, attitudes, and beliefs) of community members with regard to leprosy in Nigeria.

Objective: This study aims to assess the impact of audio-delivered and written health education on the perception of leprosy. Specific objectives are to (1) investigate the perception (local beliefs, knowledge, and attitudes) of community members toward leprosy and persons affected by leprosy; (2) investigate whether there is a difference in impact on perception between participants who have received audio-delivered health education and those who have received written health education, with specific reference to gender differences and differences between rural and urban areas; and (3) assess the impact of the participatory development of the audio-delivered and written interventions on empowerment and internalized stigma of persons affected by leprosy who developed the interventions. Additionally, we will translate and cross-culturally validate 4 study instruments measuring outcomes in 2 major Nigerian languages.

Methods: We will use a mixed methods, cross-sectional study design for the intervention development and a 3-arm cluster randomized controlled trial for its implementation and evaluation, comprising (1) baseline assessments of knowledge, attitudes, perceptions, and fears of community members, to develop the audio-delivered content and written material, and the self-esteem and internalized stigma of persons affected by leprosy; and (2) participatory development of the audio-delivered content and written material by persons affected by leprosy and the pilot and implementation of the interventions. This will be done among...
different groups (selected using cluster randomization) that will be compared (control group, audio-intervention group, and written material group) to evaluate the intervention and the impact of developing the intervention on the persons affected.

**Results:** This study was funded in June 2022, and community member participant recruitment started in January 2023. Baseline data collection was completed by May 2023 (n=811). Participatory co-creation of the audio and written health education content began in July 2023, and the materials are currently under development. Study results are expected in September 2024.

**Conclusions:** Study findings will contribute to developing evidence-based, context-specific behavioral change interventions, which are critical to addressing stigma in many leprosy-endemic communities where leprosy is highly stigmatized, and contribute toward global triple zero leprosy efforts.

**Trial Registration:** Pan African Clinical Trial Registry PACTR202205543939385; https://pactr.samrc.ac.za/TrialDisplay.aspx?TrialID=23667

**International Registered Report Identifier (IRRID):** DERR1-10.2196/53130

(JMIR Res Protoc 2024;13:e53130) doi:10.2196/53130

**KEYWORDS**

audio health education; community perception of leprosy; health education; leprosy; Nigeria; persons affected by leprosy

---

**Introduction**

**Overview**

Leprosy is an infectious disease caused by *Mycobacterium leprae* [1]. Leprosy has been stigmatized since ancient times [2]. Stigma refers to a negative social response from one group toward a low-power, stigmatized group [3]. Stigma is “a social process, experienced or anticipated, characterized by exclusion, rejection, blame, or devaluation that results from experience, perception, or reasonable anticipation of an adverse social judgment about a person or group” [4]. Stigma can occur at different levels, for example, at the intrapersonal, interpersonal, community, institutional, and structural levels [5]. Many people affected by leprosy experience the negative consequences of their condition [6,7]. Stigma can have a very negative effect on employment and education opportunities, social interaction, housing, and access to health care [8]. Stigma can also indirectly cause stress and negatively impact mental well-being, quality of life, and physical health [8]. Stigma can delay seeking treatment [9] and can also impede adherence to treatment [10]. Family members and friends may experience courtesy stigma (stigma by association) [11,12].

Perception, which refers to how individuals or groups “see” an object, person, event, or institution [13-15], is an important driver of stigma [4]. The origin of stigma lies in public perceptions about people who are stigmatized. Perception comprises knowledge, beliefs, and attitudes, which are in turn influenced by personal factors (e.g., personality and experience) and environmental factors (e.g., culture and religion) [13,14]. Leprosy-related stigma is mainly caused by fears of contagion, external manifestatons, and disabilities; religious and cultural beliefs; misconceptions; and a lack of knowledge [2,16]. Personal characteristics such as age, gender, occupation, education, and living area have also been associated with leprosy-related community stigma [17-20]. Stigma reduction (which often consists of positively influencing the perception of leprosy and increasing knowledge of leprosy) is crucial to improving the lives of people affected by leprosy and to improving leprosy services. Several interventions have successfully reduced leprosy-related stigma [21-24]. Interventions that are culture-specific and contextualized tend to be more effective [25,26].

Nigeria is among 23 global priority countries identified by the World Health Organization (WHO). In 2019, a total of 2424 new patients with leprosy were detected in Nigeria, 15% of whom had Grade 2 disabilities (G2Ds) [27]. Nigeria is among the few countries that reported more cases of G2Ds in 2019 than in previous years [27]. G2D among new patients is used as an indicator of late detection of leprosy. In addition, G2Ds (visible impairments) often exacerbate stigma and discrimination [28,29]. Leprosy is a stigmatized disease in Nigeria [30,31]. In a study among persons affected by leprosy and community members in Western Nigeria, the stigmatization of leprosy was mainly linked to perceived infectivity and perceived immoral behavior [31].

Despite progress in the control of leprosy in the last decades, the disease remains highly stigmatized [7], especially in Nigeria [30,31]. High levels of community stigma as well as internalized stigma among persons affected by leprosy may result in delayed diagnosis and disabilities [32]. To break this vicious cycle, it is important that context-specific attitude and behavior change activities are carried out. From other studies, we know that changing perceptions and improving knowledge can lead to behavior change [18,33,34]. For example, between March and June 2020, several printed materials that aimed to improve the perception of leprosy and increase knowledge about leprosy were evaluated in Uttar Pradesh, India. These context-specific materials were developed as part of NLR International’s Post-Exposure Prophylaxis project. Analysis revealed an association between the number of posters seen and a positive change in knowledge and stigma scores [35].

Access to appropriate health information is an essential step in the fight against stigma and discrimination. However, written materials are not suitable for populations with low educational levels [36]. To make health information accessible to people who are illiterate in many leprosy-endemic communities of the global south (especially women and girls), it is important to provide information through modalities other than printed materials. It is believed that audio-delivered interventions would be more suitable for people who are illiterate and low-literate.
Audio-delivered interventions, for example, radio, have been shown to be successful for different stigmatized conditions such as HIV, leprosy, and albinism [37-39].

This study aims to assess the impact of an audio-delivered intervention on the perception (knowledge, beliefs, and attitudes) of community members regarding leprosy by comparing an audio-delivered intervention with written health education. The study will be conducted in Nigeria.

Primary and Secondary Objectives

Our primary objective is to assess the impact of audio-delivered and written health education on the perception of leprosy in Nigeria. Additionally this study has the following secondary objectives: (1) to investigate the perception (local beliefs, knowledge, and attitudes) of community members toward leprosy and persons affected by leprosy; (2) to investigate whether there is a difference in impact on perception between participants who have received audio-delivered health education and those who have received written health education, with specific reference to gender differences and differences between rural and urban areas; and (3) to assess the impact of the participatory development of the audio and written interventions on empowerment and internalized stigma of persons affected by leprosy who developed the interventions.

Methods

Study Design

We will conduct a cross-sectional study (intervention development) and a 3-arm cluster randomized controlled trial (RCT; intervention implementation and evaluation). The three arms consist of (1) an intervention group who will receive the audio-delivered intervention, (2) an intervention group who will receive the written intervention (poster or flyer), and (3) a control group who will not receive any intervention.

Study Location and Setting

The study will be carried out in 6 local government areas (LGAs), 3 each in Cross River State (Boki, Calabar-South, and Obubra) and Taraba State (Jalingo, Yorro, and Zing). The 3 LGAs per area will be selected based on similarity in terms of literacy rate and prevalence of leprosy. Cross River is located in southern Nigeria, while Taraba is in the north. The total population of the study area is 940,540, estimated from the 2006 census, with 542,494 in Cross River and 398,046 in Taraba states. Both states have been selected because of the high literacy rate and prevalence of leprosy. Cross River notified a total of 106 new cases (17% with G2D) in 2019. The most common language spoken in the selected LGAs in Cross River State is Nigerian Pidgin. Other languages spoken in Cross River State include Efik, Eko, and Yala. The illiteracy rate (those with no schooling or primary education who cannot read at all) is 12.2% and 26.4% for male and female individuals, respectively. Occupations include farming, trading, and civil service employment. About 58% of women own mobile phones, compared with 71% among men [40].

Cross River State notified a total of 106 new cases (17% with G2D) in 2019. The most common language spoken in Cross River State is Nigerian Pidgin. Over 90% of the people living in the selected LGAs in southern Nigeria speak Nigerian Pidgin. Other languages spoken in Cross River State include Efik, Eko, and Yala. The illiteracy rate (those with no schooling or primary education who cannot read at all) is 12.2% and 26.4% for male and female individuals, respectively. Occupations include farming, trading, and civil service employment. About 58% of women own mobile phones, compared with 71% among men [40]. Cross River State’s poverty index is 0.146, and the poverty headcount rate is 36.3%. The case notification rate for leprosy is 2.76 per 100,000 people based on 2020 data.

Taraba is among the 15 high-burden states for leprosy in Nigeria; an average of 100 new cases were detected between 2013 and 2017, with an average of 6% and 5% of children and G2D cases, respectively. This figure is likely to be underreported, considering the significant presence of isolated populations such as nomadic pastoralists and internally displaced people in the state.

The most common language spoken in the selected LGAs in Taraba State is Hausa; 80% to 90% of the people living in the selected LGAs in Northern Nigeria speak Hausa. Other languages spoken in Taraba State include Mummuye and Fulfulde. The illiteracy rate is 30.1% and 64.9% in Taraba State for male and female individuals, respectively. Occupations in Taraba include farming, mainly crop production and cattle rearing, petty trading, and civil service employment. The majority of men (71%) own a mobile phone, while only 44.6% of women do. Taraba State’s poverty index is 0.448, and the poverty headcount rate is 87.7%. The case notification rate for leprosy is 2.30 per 100,000 people.

Study Population

The following 2 groups of participants will be included in the study: persons affected by leprosy (for the participatory development of the interventions) and community members (the target group of the interventions).

Inclusion and Exclusion Criteria

Individuals aged 18 years or older will be included in this study. Individuals who do not speak Nigerian Pidgin or Hausa and who are unable or unwilling to give informed consent will be excluded from the study.

Study Duration and Sample Size Calculation

This study’s duration will be 2 years. The sample size for the various components of the study is as follows. (1) A total of 200 community members (100 for each language per study area) will be included in the cross-cultural validation of the Explanatory Model Interview Catalogue Community Stigma Scale (EMIC-CSS) and Social Distance Scale (SDS), while 100 persons affected by leprosy (50 for each language per study area) will be included in the cross-cultural validation of the Internalized Stigma of Mental Illness (ISMI) scale and Rosenberg Self-Esteem Scale (RSES; see below). (2) A total of 770 community members will be included in the baseline and follow-up questionnaire interviews. This means a random sample of at least 385 persons in Taraba State (northern Nigeria) and at least 385 persons in Cross River State (southern Nigeria), which will consist of 114 in the audio intervention group, 114 in the written material intervention group, and 157 in the control group in each region.

This RCT will have three arms: (1) an audio-delivered intervention group, (2) a written material intervention group, and (3) a control group. The sample size calculation is based on 2 calculations. The intervention group calculation is based on an estimate of the difference in knowledge improvement...
between the audio-delivered intervention and the written material intervention groups. We used data from a perception study in India. In this study, postintervention scores improved by 12.5% after a poster intervention and community meetings. We estimate that the effect of posters alone would be an increase of 10%. We want to be able to detect an improvement of at least 15% between the audio-delivered and the written material intervention groups. The sample size of the intervention group is therefore based on a proportion 1 of 10 (estimated percentage of improvement in knowledge of leprosy in the written intervention group) and a proportion 2 of 25 (ie, an improvement of 15% or more). With a power of 80%, a significance level of .05, and a 15% loss to follow-up, 114 participants are needed in each intervention group.

We expect an increase of 2% in the knowledge score in the control group. The control group calculation is therefore as follows: proportion 1:10, proportion 2:2, power 80%, significance .05, and 15% loss to follow-up, resulting in 157 participants.

A total of 25 people will be included in the participatory development of the material [41]. In each state, these will consist of 10 persons affected by leprosy and 2-3 community members. Semistructured interviews will be conducted until data saturation is reached.

**Description of the Intervention**

The interventions consist of (1) an audio-delivered intervention and (2) a written or printed intervention (such as posters or flyers) for education on leprosy, awareness-raising, and stigma reduction.

We will compare the effect of the interventions with a control group. The audio-delivered and written content will be developed based on local beliefs, misconceptions, and fears about leprosy identified in the baseline study. This will be done using participatory approaches. A group of persons affected by leprosy and a few members of the community will be formed, who will be guided by a researcher to develop the messages and materials (participatory development). The key messages of the audio-delivered and written interventions will be the same. The materials will be developed in the main languages spoken in the study areas: Nigerian Pidgin (Cross River State) and Hausa (Taraba State). The majority (>80%) of our target group speaks the language used for this study. The audio-delivered intervention will be incorporated into Audiopedia [42]. Audiopedia’s website was designed to provide access to open knowledge foremost on health, livelihood, and well-being to both community-based organizations and nongovernmental organizations and individuals. Community-based organizations and nongovernmental organizations can benefit from using Audiopedia as part of their social and behavior change communication strategy, as it enables them to search, download, embed, and share audio files. Approximately 5000 audio clips, with a total runtime of 150 hours, are available in 11 languages.

Audiopedia was optimized for search engines, thus making contents easy to find. Audiopedia also provides several technological solutions to make audio-based contents accessible to both literate and illiterate audiences, such as solar-powered audio players, mobile web applications for smart (feature) phones, and Wi-Fi hotspots that can stream audio-based contents without the need for internet connectivity, etc.

**Participant Recruitment and Follow-Up**

Persons affected by leprosy and community members will be selected based on purposive sampling. They will develop the content of the interventions. This is therefore, ideally, a diverse group of people who represent multiple perspectives. We will select participants based on purposive sampling to ensure adequate representation of age, gender, and villages.

The intervention will be implemented in North and South Nigeria, 2 areas that are very different. Therefore, we will cluster-randomize the interventions to ensure comparable groups are included in both areas. There are three different “groups” in each site (North and South): (1) the audio-delivered intervention, (2) the written materials intervention, and (3) no intervention (control group).

We will select 3 LGAs in North and South Nigeria (6 in total); each LGA will have either the audio-delivered or the written intervention or be a control group. The three LGAs per area will be selected based on their similarity in terms of literacy rate and endemicity of leprosy.

The interventions will be randomly allocated to clusters (LGAs) based on a random numbers list. This is a 2-stage random sampling. We will select a random sample of LGAs and a random sample of participants within each LGA. The participants in each LGA will be selected using the “spin the bottle” approach: a bottle will be spun in front of the most central place in the village (rural area) or neighborhood (urban area); the direction the bottle points at is the direction we start walking and counting. The first house to be included is selected by casting a die. Participant selection will be done by proportional sampling among the clusters. We will include the same participants at baseline and follow-up (a paired sample). This study protocol is reported in accordance with the SPIRIT statement (see checklist in Multimedia Appendix 1). The timeline schedule for participants during the study period is represented in Table 1, and a schematic description of the study flow chart is depicted in Figure 1.
Table 1. Study schedule for enrollment, intervention, and assessment.

| Schedule | Study period and time point | Close  
|----------|-----------------------------|------
|          |   Enrollment | Allocation | Postallocation | out |
|          | t₁              | t₀ (month 0) | t₁ (baseline) | t₂ (intervention) | t₃ (6 months follow-up) | t₄  |

**Enrollment**
- Eligibility screening ✓
- Informed consent ✓
- Randomization and allocation ✓

**Intervention group**
- Audio-delivered intervention ✓
- Written ✓
- Control (no intervention)

**Assessments**

**Baseline variables**
- KAP<sup>a</sup>, EMIC-CSS<sup>b</sup>, and SDS<sup>c</sup> ✓
- RSES<sup>d</sup> and L-ISMI<sup>e</sup> scale ✓
- CNA<sup>f</sup> ✓

**Outcome variables**
- KAP, EMIC-CSS, and SDS ✓
- RSES and L-ISMI scale ✓

**Analysis**
- Data analysis ✓

---

<sup>a</sup>KAP: knowledge, attitudes, and practices.
<sup>b</sup>EMIC-CSS: Explanatory Model Interview Catalogue Community Stigma Scale.
<sup>c</sup>SDS: Social Distance Scale.
<sup>d</sup>RSES: Rosenberg Self-Esteem Scale.
<sup>e</sup>L-ISMI: Leprosy-adapted Internalized Stigma of Mental Illness.
<sup>f</sup>CNA: Communication Needs Assessment.
Primary and Secondary End Points

The outputs of this study will be an audio-delivered intervention and written or printed health education material for education on leprosy, awareness-raising, and stigma reduction.

In addition, all 4 scales will be translated and cross-culturally validated in Nigerian Pidgin and Hausa languages as part of this study. The qualitative outcomes are knowledge and attitudes toward (persons affected by) leprosy (eg, local beliefs, fears, and misconceptions).

Outcome measures used to be assessed at baseline and follow-up are as follows:

- Demographic information: to disaggregate data by gender, the “knowledge, attitudes, and practices” (KAP) measure includes a form to collect demographic information.
- The KAP measure, as used in a study in India [16], covers 8 main topics: early symptoms, cause, mode of transmission, treatment, prevention, curability, contagiousness when on treatment, and prevention of disabilities. The KAP is a questionnaire and has not been formally validated as a scale (nor will it be used as a scale). We will translate and pilot test the KAP measure.
- Community stigma, using the EMIC-CSS (used in Nigeria but not validated in Nigerian Pidgin and Hausa yet [17]).
- Desired social distance toward persons affected by leprosy as a proxy for attitudes and fear, using the SDS (used among persons affected by leprosy in Nigeria but not validated in Nigerian Pidgin and Hausa yet [17]).
- Self-esteem and internalized stigma of persons affected by leprosy, using the RSES (used in Nigeria but not validated in Nigerian Pidgin and Hausa yet) and the leprosy-adapted ISMI (used among persons affected by leprosy and used in Nigeria but not formally validated yet [17]).

Data Collection

Overview

We will use a mixed methods approach and collect qualitative data (in-depth interviews and focus group discussions [FGDs]) and quantitative data (the KAP measure, EMIC-CSS, SDS, communication needs assessment, RSES, and ISMI). We will also collect demographic information (including literacy levels) from each participant.

This study consists of the steps and phases outlined in the following subsections.

Step 1

None of the tools listed in step 2 have been validated in Nigerian Pidgin and Hausa yet; therefore, they will be cross-culturally validated before use. We will assess conceptual, item, semantic, operational, and measurement equivalence using a framework for cross-cultural equivalence testing based on the work of Herrman et al [43], Terwee et al [44], and Stevelink and van Brakel [45]. In addition, the interview and group discussion guides (outlined in step 2) will be pilot tested among a small sample of participants before use. The KAP measure will be translated and pilot-tested.

Step 2

A baseline study to assess the perceptions of community members using mixed methods will be conducted, consisting of both in-depth interviews and FGD sessions and questionnaires (KAP, EMIC-CSS, and SDS). A communication needs assessment will be conducted as part of the baseline study to determine the most appropriate mode of delivery of the audio intervention and the most appropriate duration and frequency of the audio intervention. We will also conduct a baseline study among the persons affected by leprosy group who will be involved in the development of the interventions. Self-esteem and internalized stigma will be assessed using in-depth interviews, RSES, and ISMI.
Step 3
The participatory development of the audio-delivered and written content of the interventions is based on the knowledge gaps, beliefs, misconceptions, fears, and community attitudes identified in the baseline study. We will use the “6 steps in quality intervention development (6SQuID) framework [46], consisting of (1) defining and understanding the problem and its causes, (2) identifying which causal or contextual factors are modifiable: which have the greatest scope for change and who would benefit most, (3) deciding on the mechanisms of change, (4) clarifying how these will be delivered, (5) testing and adapting the intervention, and (6) collecting sufficient evidence of effectiveness to proceed to a rigorous evaluation. The final step of 6SQuID (step 6) is part of step 5 of this study—evaluation. The group of persons affected by leprosy who play a leading role in the development of the messages and materials (participatory development) will codetermine both the content of the materials (based on the knowledge gaps, beliefs, misconceptions, fears, and community attitudes identified in the baseline study) and the mode of delivery (based on the most appropriate means of communication determined by the communication needs assessment conducted as part of the baseline study).

Step 4
Implementation of the interventions in the study areas will be done in 2 groups: one intervention group will receive the audio-delivered intervention, and the other intervention group will receive the written intervention (poster or flyer). A control group will not receive any intervention. The groups will be cluster-randomized.

Step 5
Step 5 involves evaluating the impact of (1) the intervention on the community and (2) developing the intervention on persons affected, using the same mixed methods as the baseline studies (described in step 2). It should be noted that the in-depth interviews at baseline are mainly conducted to get insight into specific fears, local beliefs, and misconceptions about leprosy among community members, as well as insight into the self-esteem and internalized stigma of persons affected by leprosy. At follow-up, the communication needs questionnaire will not be administered anymore. In addition, FGDs will be held with community members at baseline and follow-up in the 6 LGAs and with the persons affected by leprosy group. In addition to further exploring themes that arose during the in-depth interviews, additional questions will be asked (at follow-up) to get insight into, among other things, awareness of, experiences with, and exposure or access to the audio and written interventions; thoughts about content; mode of delivery and frequency; and strengths and points of improvement.

Data Management and Data Analysis
Quantitative data collection will be done using electronic forms developed in the Open Data Kit, and all data will be securely stored in the cloud with access only to the research team. The recordings of the in-depth interviews and FGDS will be transcribed to the local languages, translated to English, and analyzed by 2 independent researchers using open, inductive coding and content analysis. Similar phrases with recurring themes will be coded in NVivo (QSR International). Quantitative data will be collected in the KoboCollect (Kobo Inc) mobile phone app. Data analysis will be done in the software package SPSS Statistics (SPSS Inc). Simple descriptive methods will be used to generate a demographic profile of the study sample. Differences between participants in the groups (audio, written materials, and control groups) will be evaluated using the Mann-Whitney U test or 1-tailed t test for continuous variables and the chi-square statistic for categorical variables. The mean (SD) or median (IQR), depending on the distribution of the data, of the total scores of the scales used will be calculated per intervention area. Stepwise multivariate regression with backward elimination will be done to examine what factors will have an independent effect on the outcomes. We will calculate the percentage change and corresponding 95% CI before and after the interventions are implemented and the statistical significance of this difference using a z test for differences between proportions. Effect sizes will also be calculated. If necessary, we will correct for differences in demographic information between study arms using quantile regression. We will compare the differences in follow-up assessment between the audio-delivered and written interventions. The knowledge gaps, beliefs, misconceptions, fears, etc identified in the baseline study will be summarized into a narrative review and will be used by the persons affected by leprosy group as input for the development of the content of the interventions. The most appropriate modes of delivery will be determined based on the communication needs assessment that will be conducted as part of the baseline study. The interventions will be developed using participatory methods.

Confidentiality and anonymity of data will be ensured in data collection, storage, analysis, and publication. Research assistants who will collect the data will be trained in data management, the maintenance of confidentiality, and ensuring privacy during data collection. Data will only be analyzed and shared with the Dutch and German researchers (outside of Nigeria) when they have been fully anonymized. The lead applicant will take full responsibility for ensuring the appropriate storage and security of data. Data will be kept for 5 years and destroyed after this time frame when no longer required.

Ethical Considerations
Ethical approval has been obtained from the Health Research and Ethics Committee of the University of Nigeria Teaching Hospital, Ituku-Ozalla, Enugu (NHREC/05/01/2008B-FWA00002458-1RB00002323). In addition, appropriate clearance was sought and obtained from the respective ethics committees of both the Taraba and Cross Rivers State Ministries of Health. The confidentiality of the study participants and the data collected from them will be ensured.

Written informed consent (Multimedia Appendix 2) will be obtained from the study participants. The questionnaire will be translated into the local languages included in this study and back-translated to ensure the soundness of the translation process. The consent form explicitly states the right of the participant to refuse giving consent or withdraw from the study.
at any point. Also, he or she can decline to answer any question. Informed consent will be obtained from each participant before participation. Each participant will be given a copy of the participant information sheet and consent form to keep.

This RCT study was prospectively registered with and listed on the Pan African Clinical Trials Registry (PACTR202205543939385).

Results

This study was funded in June 2022, and community member participant recruitment started in January 2023. Baseline data collection was completed by May 2023, with a total of 811 respondents. Participatory cocreation of the audio and written health education content began in July 2023, and the materials are currently under development. Intervention will be administered after pilot-testing, and the follow-up period will last for 6 months. Neither the implementation of the study intervention nor the data analysis have commenced as of the time of submission. Study results are expected in September 2024.

Discussion

Principal Considerations

In this study, we will apply mixed methods to compare the impact of audio-delivered versus written health education intervention materials and a control group without any intervention on the perception (knowledge, beliefs, and attitudes) of community members regarding leprosy. Through a participatory cocreation process, we will develop the content of the intervention materials alongside persons affected by leprosy and implement a 3-arm cluster RCT to evaluate the effect of the intervention. The effect of the cocreation process on internalized stigma among the affected persons will also be evaluated.

Developing context-specific behavioral change intervention activities is critical to addressing stigma in many leprosy-endemic communities where leprosy is highly stigmatized. The outcome of this intervention is expected to influence knowledge and, hopefully, improve the attitudes of community members toward leprosy, as well as improve self-esteem among persons affected by leprosy. In the long run, early leprosy case finding will be boosted following stigma reduction. To the best of our knowledge, this is the first study of this kind in Nigeria. The generated educational materials (both the audio and written content) are reusable products that may be adopted or used by the national program. The audio-delivered educational content will be freely available on the Audiopedia platform, licensed under Creative Commons (CC BY).

Limitations

There are inherent limitations associated with cluster-randomized studies, such as cluster size variability. With the challenge of achieving uniform cluster sizes, variability in cluster size could impact the generalizability of the results. Additionally, there is a risk of contamination between clusters; however, this is minimal given the geographic distance between the study cluster locations.

The study sites were purposefully selected due to their similarity in having a high prevalence of leprosy and G2Ds; thus, the findings from this study may not be generalizable to the entire Nigerian populace.

Conclusion

Outputs from the research will offer policy makers, national and regional program managers, and partners reliable evidence for a new approach toward stigma reduction activities, thereby contributing toward global triple zero leprosy efforts. Content generated provides pragmatic and contextual evidence-based tools for effective health education campaigns and awareness creation, especially among populations with low literacy levels, both in Nigeria and other low- and middle-income countries.

Acknowledgments

This work is supported by the Leprosy Research Initiative Foundation (grant FP_22/15). The funding source had no role in the study design, collection, analysis, and interpretation of data; the writing of the report; or the decision to submit the article for publication.

The authors deeply appreciate the contributions of persons affected by leprosy and the kind support of the central unit of the National Tuberculosis, Leprosy, and Buruli Ulcer Control Programme, Federal Ministry of Health, Abuja, Nigeria, as well as the respective team members of both Cross River State and Taraba State Tuberculosis, Leprosy, and Buruli Ulcer Control Programme.

Data Availability

This is a study protocol manuscript. Deidentified research data that support the findings from this study will be made publicly available when the study is completed and published.

Authors' Contributions

AVTN and CG conceived the original research idea. AVTN, NMO, and JC contributed to fully conceptualizing the research proposal. AVTN, NMO, JC, CN, TD, SA, CG, UAO, AM, CE, OE, and NE all contributed to the design and planning of the study. NMO wrote the draft of the study protocol manuscript and the SPIRIT (Standard Protocol Items: Recommendations for
Interventional Trials) checklist. SA illustrated the study flow chart. All authors discussed the study protocol and contributed to the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents.
[DOCX File, 45 KB - resprot_v13i1e53130_app1.docx ]

Multimedia Appendix 2
Participant information sheet and informed consent form.
[DOCX File, 29 KB - resprot_v13i1e53130_app2.docx ]

Multimedia Appendix 3
Peer review report by Leprosy Research Initiative (LRI), Amsterdam, The Netherlands.
[PDF File (Adobe PDF File), 117 KB - resprot_v13i1e53130_app3.pdf ]

References


30. Arulanantham S. Addressing inequality and exclusion—the opinion of people affected by leprosy in Africa and Asia, as to what should be included in any post Millennium Development Goal framework. Lepr Rev 2014;85(3):133-140 [FREE Full text] [doi: 10.47276/hr.85.3.133]


42. Audiopedia. URL: https://www.audiopedia.io [accessed 2021-01-21]


Abbreviations

6SQuID: 6 steps in quality intervention development
EMIC-CSS: Explanatory Model Interview Catalogue Community Stigma Scale
FGD: focus group discussion
G2D: grade 2 disability
ISMI: Internalized Stigma of Mental Illness scale
KAP: knowledge, attitudes, and practices
LGA: local government area
RCT: randomized controlled trial
RSES: Rosenberg Self-Esteem Scale
SDS: Social Distance Scale
WHO: World Health Organization

©Ngozi Murphy-Okpala, Tahir Dahiru, Anna T van ’t Noordende, Carolin Gunesch, Joseph Chukwu, Charles Nwafor, Suleiman Hudu Abdullahi, Chukwuma Anyaikê, Ugochinyere Angelic Okereke, Anthony Meka, Chinwe Eze, Okechukwu Ezeakile, Ngozi Ekeke. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.01.2024. 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.
Adaptive Intervention to Prevent Respiratory Illness in Cerebral Palsy: Protocol for a Feasibility Pilot Randomized Controlled Trial

Alyssa Fleischman, MD; Carlos Lerner, MD, MPhil; Heidi Kloster, MD; Paul Chung, MD, MS; Thomas Klitzner, MD, PhD; Christopher Cushing, PhD; Danielle Gerber, BA; Barbara Katz, MA; Gemma Warner, MSSW; Kristina Devi Singh-Verdeflor, MPH; Roxana Delgado-Martinez, BA; Lorena Porras-Javier, MPH; Siem Ia, CPNP-PC; Teresa Wagner, RN; Mary Ehlenbach, MD; Ryan Coller, MD, MPH

Department of Pediatrics, University of Wisconsin School of Medicine and Public Health, Madison, WI, United States
Department of Pediatrics, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States
Department of Health Systems Science, Kaiser Permanente School of Medicine, Pasadena, CA, United States
Department of Health Systems Science, Kaiser Permanente School of Medicine, Pasadena, CA, United States
Clinical Child Psychology Program and Schiefelbusch Life Span Institute, University of Kansas, Kansas, KS, United States
Family Voices of Wisconsin, Madison, WI, United States
UW Health Kids American Family Children's Hospital, Madison, WI, United States

Corresponding Author:
Alyssa Fleischman, MD
Department of Pediatrics
University of Wisconsin School of Medicine and Public Health
H6/551 CSC
600 Highland Ave
Madison, WI, 53792
United States
Phone: 1 608 262 9150
Email: afleischman2@wisc.edu

Abstract

Background: This study will pilot-test an innovative just-in-time adaptive intervention to reduce severe respiratory illness among children with severe cerebral palsy (CP). Our intervention program, Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT), delivers timely customized action planning and rapid clinical response when hospitalization risk is elevated.

Objective: This study aims to establish RE-PACT’s feasibility, acceptability, and fidelity in up to 90 children with severe CP. An additional aim is to preliminarily estimate RE-PACT’s effect size.

Methods: The study will recruit up to 90 caregivers of children with severe CP aged 0 to 17 years who are cared for by a respiratory specialist or are receiving daily respiratory treatments. Participants will be recruited from pediatric complex care programs at the University of Wisconsin–Madison (UW) and the University of California, Los Angeles (UCLA). Study participants will be randomly assigned to receive usual care through the complex care clinical program at UW or UCLA or the study intervention, RE-PACT. The intervention involves action planning, rapid clinical response to prevent and manage respiratory illness, and weekly SMS text messaging surveillance of caregiver confidence for their child to avoid hospitalization. RE-PACT will be run through 3 successively larger 6-month trial waves, allowing ongoing protocol refinement according to prespecified definitions of success for measures of feasibility, acceptability, and fidelity. The feasibility measures include recruitment and completion rates as well as intervention satisfaction. The fidelity measures include observed versus expected rates of intervention and data collection activities. The primary clinical outcome is a severe respiratory illness, defined as a respiratory diagnosis requiring hospitalization. The secondary clinical outcomes include hospital days and emergency department visits, systemic steroid courses, systemic antibiotic courses, and death from severe respiratory illness.

Results: The recruitment of the first wave began on April 27, 2022. To date, we have enrolled 30 (33%) out of 90 participants, as projected. The final wave of recruitment will end by October 31, 2023, and the final participant will complete the study by April 30, 2024. We will start analyzing the complete responses by April 30, 2024, and the publication of results is expected at the end of 2024.
Conclusions: This pilot intervention, using adaptive just-in-time strategies, represents a novel approach to reducing the incidence of significant respiratory illness for children with severe CP. This protocol may be helpful to other researchers and health care providers caring for patients at high risk for acute severe illness exacerbations.

Trial Registration: ClinicalTrials.gov NCT05292365; https://clinicaltrials.gov/study/NCT05292365

International Registered Report Identifier (IRRID): DERR1-10.2196/49705

(JMIR Res Protoc 2024;13:e49705) doi:10.2196/49705

KEYWORDS
just-in-time adaptive intervention; respiratory illness; cerebral palsy; action planning; digital health

Introduction

Background

Children with severe cerebral palsy (CP) have spastic quadriplegia and are classified in level IV or V on the Gross Motor Function Classification System (GMFCS), often resulting in little or no independent mobility and serious respiratory consequences [1]. The mechanisms of respiratory illness in severe CP vary, parallelising those of other neuromuscular diseases [2], and include respiratory muscle weakness, recurrent infections and aspiration with inflammatory fibrosis, impaired airway clearance from altered tone, upper airway abnormalities, and poor chest wall compliance [3,4].

Respiratory illness is consistently the leading cause of death and hospitalization in severe CP [5,6]. Respiratory illness accounts for 59% of the deaths [5,7] and 25% of the hospitalizations [8-10] in severe CP. Moreover, respiratory illness strongly predicts future risk: respiratory hospitalization risk is 10-fold higher with a respiratory illness in the past year. Nevertheless, respiratory illness risk factors in severe CP are considered modifiable [11]. The prevention of these events is a significant need and a key to improving the quality of life and decreasing mortality [1,12].

Preventing hospitalization requires the opportunity for families and clinical teams to connect early enough to change trajectory [13-15]. Parents of children with CP have expressed the need for interventions focused on crisis management and self-efficacy [9,13,16]. However, respiratory illness in severe CP has broad comorbid triggers (eg, emesis, dysphagia, aspiration, and seizures). Because of this complexity, simple action plans or coaching alone may not address the breadth of respiratory illness triggers or potential responses; for example, if a parent of a child with severe CP follows an action plan directed toward bronchospasm, it would not effectively address an acute infectious lower respiratory infection. Parents of children with severe CP need comprehensive action planning and coaching; they also need an efficient direct extension to their clinical team for a just-in-time (JIT) adaptive clinical response directed specifically to acute real-time problems.

Currently, difficulty identifying when JIT care is needed is a barrier to effective illness response. Concerns may not reach clinical teams until an emergency department (ED) visit or hospitalization is inevitable. A national expert panel to identify interventions to prevent the hospitalization of children with complex diseases concluded that enhanced access, proactive crisis planning, and support for caregiver technical skills were crucial strategies to lower hospital use [17]. Prior postdischarge research has confirmed that admissions and ED visits could be better predicted by identifying when parents were not confident that their child with chronic conditions could avoid hospitalization or an ED visit than by other clinical or demographic indicators [18,19]. Preliminary work with a cohort that included children with severe CP demonstrated that parent confidence, monitored prospectively and repeatedly by SMS text message, is feasible, is acceptable, and predicts hospitalization within 2 weeks. This program of research will drive care forward by providing JIT care triggered by parents’ self-reported period of low confidence, thus matching the intervention to the immediate clinical need and preventing respiratory crisis.

Prior Work

This team developed the earlier Plans for Action and Care Transitions (PACT) intervention to prevent hospitalizations for children with complex chronic diseases, including severe CP. After integrating a systematic literature review [20], parent interviews [21], and a national expert panel [17], each focused on preventing hospitalization, the team designed PACT to leverage evidence-based strategies from different populations: asthma action planning [22-24], health coaching [25-27], and feedback from parent advisory group meetings. The PACT intervention delivered action planning and coaching activities to children with diverse complex diseases, including severe CP, and observed 40% lower hospitalization rates for intervention versus control patients [28]. Simultaneously, our prior multisite research observed that confidence to avoid hospitalization, measured through repeated SMS text messaging, predicted hospitalization over the subsequent 2 weeks. Our clinical team and family partners hypothesized that periods of low confidence might be a useful tailoring variable to prompt intervention delivery [29].

The PACT intervention has now been adapted to prevent severe respiratory illness in children with severe CP and to integrate SMS text messaging as a tailoring variable within a JIT adaptive intervention framework [29,30].

Objectives

This pilot study (ClinicalTrials.gov: NCT05292365) is designed to establish the feasibility, acceptability, and fidelity of our intervention program, Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT) in up to 90 children with severe CP and to establish a preliminary effect size of RE-PACT to inform a future efficacy study to reduce severe
respiratory illness. This intervention consists of three related parts: (1) universal action planning, (2) an ongoing assessment of hospitalization risk, and (3) an algorithm to determine when to increase clinician contacts and tailor action plans. The study period will be divided into 3 waves; after each wave, feasibility, acceptability, and fidelity data will be reviewed against predefined measures of success to adjust the protocol and overcome implementation barriers. We describe the design and protocol of this trial in the following sections.

Methods

Participants and Setting
This intervention will recruit primary caregivers of children with severe CP. Up to 90 caregivers of children with severe (GMFCS level IV or V) CP aged 0 to 17 years and cared for by a respiratory specialist or receiving daily respiratory treatments will be enrolled. Participants will be recruited from pediatric complex care programs at the University of Wisconsin–Madison (UW) and the University of California, Los Angeles (UCLA). These programs were established to deliver care to children with medical complexity. The key components of each program include pediatric clinicians, care coordinators, and extended visit lengths, which aid in delivering comprehensive care to children with CP. Both clinical programs have been described in more detail elsewhere [28,31].

Inclusion Criteria
Participants are caregivers of children with severe CP. Individuals must meet all inclusion criteria to be eligible to participate in the study. Caregiver criteria include (1) being aged at least 18 years, (2) being the primary caregiver to an eligible child, (3) ability to speak English or Spanish well enough to be interviewed, and (4) having a mobile phone capable of sending and receiving SMS text messages. Child criteria include (1) age 0 to 17 years, (2) GMFCS level IV or V CP [32], and (3) being cared for by a respiratory specialist or receiving daily respiratory treatments (oxygen, ventilation, airway clearance device, and medications).

Exclusion Criteria
During this study, participants are asked to reply to SMS text messages when received at random times during daytime hours and connect with an intervention clinical responder either at home, in person at a mutually agreeable location, by mobile phone, or over the internet. Any individual lacking the ability or willingness to engage in SMS text messaging or clinical responder interactions during the study will be excluded from participation in the study.

Recruitment and Screening
We will recruit caregivers of children with severe CP aged between 0 and 17 years. We will recruit up to 90 participants (n=45, 50% at each site) divided across 3 waves. In each wave, there is a 1- to 2-month enrollment period. We anticipate that approximately 80% of those screened will enroll, requiring approximately 110 individuals to be screened.

Using diagnostic codes for CP (International Classification of Diseases, Tenth Revision [ICD-10]: G80-83), we will identify potential participants by reviewing clinic registries and electronic health record data, which contain detailed information about children and their diagnoses. We will send an opt-out letter that alerts families that a research study is being conducted and their child may be eligible, with a contact number to call if they wish to opt out of the research or if they wish to receive additional information or have any questions. Potentially eligible caregivers will be contacted by telephone to screen for eligibility and interest.

If the research team is not notified that a family wishes to opt out of the research, the study research personnel will attempt to call the families (or meet them at an upcoming visit) to complete screening, informed consent, baseline questionnaires, and random group assignment. CP status and additional eligibility criteria will be determined with a reliable and valid parent questionnaire and screener conducted at the beginning of the initial telephone contact [32].

Individuals who do not meet the criteria for participation in this trial (screen failure) because they meet ≥1 exclusion criteria that are likely to change over time may be rescreened. Theoretical examples might include a child developing a need for respiratory treatment or families acquiring a mobile phone capable of sending and receiving SMS text messages.

All study participants will undergo informed consent, including authorization to view the child’s medical record and participate in action planning, rapid clinical response, and weekly SMS text message surveillance.

Study Design
This is a 2-site pilot randomized controlled clinical trial to establish the RE-PACT protocol’s feasibility, acceptability, and fidelity as well as an estimate of effect size. We anticipate being underpowered to assess the efficacy of the intervention in this pilot study; however, to inform future randomized controlled trial power estimates, we will test differences between the intervention and control groups in primary and secondary clinical outcomes.

Study participants will be randomly assigned to receive usual care through the complex care clinical program at UW or UCLA or the study intervention, RE-PACT. Random allocation will be concealed from the research staff conducting recruitment and will use a 1:1 allocation with random block sizes of 2 and 4. Block randomization will be achieved with a computer-generated random number list prepared by the study biostatistician without clinical involvement in the trial. Randomization will be stratified by site to account for site-specific study characteristics.

RE-PACT will be run through 3 successively larger 6-month trials (waves), allowing ongoing protocol refinement between waves, guided by prespecified definitions of success for feasibility, acceptability, and fidelity measures. Each wave has a specific protocol refinement focus (wave 1: onboarding, training, recruitment, and data collection; wave 2: randomization and intervention activities; and wave 3: rapid enrollment and the conduct of all protocol activities with high fidelity). Participants in both groups will undergo assessments of demographic, clinical, and caregiving measures using
questionnaires and medical record review case report forms at baseline and at 6 months after enrollment. Intervention feasibility, acceptability, and fidelity data will be collected from parent reports, medical records, and research team logs using case report forms.

**Description of the Intervention**

RE-PACT uses a dynamic JIT adaptive intervention design [33] to deliver proactive intervention based on risk modeling and partnership between the care team, patients, and families. Although the causes of respiratory illness in severe CP are modifiable, they are also broad and require distinct responses, even for the same child, over time. RE-PACT assumes that (1) every patient with severe CP has a risk of hospitalization, (2) some risks are knowable via the ecosystem of data generated around patient care, and (3) an intervention delivered when risk is increasing can reduce hospitalizations. RE-PACT’s design addresses the changing needs of a child and family. RE-PACT involves action planning, rapid clinical response to prevent and manage respiratory illness, and weekly SMS text messaging surveillance of caregiver confidence for their child to avoid hospitalization.

**Action Planning**

**Overview**

All intervention families will receive respiratory illness action plans within 1 month of study entry. The action plan format and process are adapted from the original PACT study, and the contents include (at minimum) recognizing, describing, and managing the child’s known contributors to respiratory illness. The three main components of the action plans are (1) **focus area** for the action plan (eg, asthma, aspiration, and seizures); (2) **severity levels** corresponding to objective and subjective indicators of baseline (green), concerning (yellow), and severe (red) statuses (eg, >2 L/min of oxygen); and (3) **specific actions** that caregivers should take to manage each status (eg, increase vest therapy, albuterol, suction every 4 hours, and use oxygen up to 4 L/min). As needed, JIT plans are also created at times of low confidence by parent request or by clinician determination during the study period. Any plan created will be developed with families, target an issue that plausibly will recur and lead to respiratory illness–related ED or hospital visit, and, when relevant, be harmonized with prior plans and reflect pulmonologist agreement.

**Mobile Health Platform**

The mobile health (mHealth) platform is built from an earlier study, *Assessing Confidence at Times of Increased Vulnerability (ACTIV)* [29], which was designed to elicit a SMS text rating of confidence to avoid hospitalization in the next month (ratings range from 1 to 10, where 1 is lowest confidence, and 10 is highest confidence; Figure 1). The platform supports English and Spanish languages. Beginning on the Sunday after enrollment, families will start receiving weekly SMS text messages asking them to rate their confidence for their child to avoid hospitalization in the next month. SMS text messages are programmed to be sent at random days and times to caregivers, averaging once weekly (Sunday to Thursday) between 8 AM and 9 PM (local time). The Sunday-to-Thursday time frame was chosen to support a feasible response during business days and hours. After 2 hours of nonresponse, a reminder is sent, and this is repeated up to 2 times at 2-hour intervals. Clinical responders will receive an email notification in real time if a participant reports low confidence. In addition, clinical responders will receive an SMS text message notification between 9 AM and 6 PM with the report of low confidence. If a response comes outside of these hours, it will be delayed until the next day.

**Figure 1.** Schematic of the mobile health (mHealth) SMS text message process in which all study participants receive weekly SMS text messages asking them to rate their confidence for their child to avoid hospitalization in the next month.
Rapid Clinical Response

This response is adapted from our prior intervention (PACT) [28]. In RE-PACT, a clinical responder guides the JIT response, adapting to the current child and family situation. Triggers for the clinical responder include (1) low family-reported confidence (a confidence rating of <5) during mHealth messaging, (2) hospital discharge, and (3) family call or electronic message to the clinic owing to acute respiratory concerns (Figure 2). Clinical responders are clinicians, including medical doctors, nurse practitioners, registered nurses, and care coordinators (or equivalent). The same responder intends to work with the family throughout study enrollment.

Figure 2. Summary of the Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT) intervention. The figure illustrates low-confidence SMS text messages as the trigger of rapid clinical response. Other triggers include hospital discharge or family-expressed respiratory concerns through telephone call or electronic message to the clinic.

Rapid clinical responses include 3 interactions between family participants and clinical responders. First, a triage contact occurs within 24 hours of a trigger (during business hours). The triage contact goals are to determine the nature of the trigger, whether an action plan exists for the situation, and whether the issue is within the clinical responder’s clinical practice scope. If not within the scope, the issue is referred to the relevant support (eg, a specialty physician or clinic social worker). Second, a response planning visit occurs either as a component of the triage contact or at a mutually agreed upon time within 72 hours of the trigger. Third, at least 2 follow-up contacts occur within 2 weeks of the trigger, with additional follow-ups as indicated by ongoing need until the issue is resolved. All contacts can occur through any of the following, at the preference of the family: telephone call, clinical encounter (telehealth, clinic, and hospitalization), or a home visit. The follow-up contacts can occur through electronic communication if the clinical responder and family determine this to be appropriate. At each contact point, there are two goals: (1) ensuring that the family understands red flags, relevant medications, and whom to call and when and keeps notes about the issue; and (2) coaching and skill transfer for the family to generate solutions and lead actions, with the responder intervening if the family is stuck or if clinical needs dictate intervention. Each contact point has scripting to guide the clinical responder as well as electronic health record documentation templates. At the end of a clinical responder event, the responder determines whether the issue affects respiratory health, is likely to recur, and poses a risk for future ED or hospital visits. If all of these are true, the responder either updates existing action plans or creates a new one to address the issue. A participant is considered to have completed the study if they have completed the baseline and 6-month follow-up assessments (Table 1; Figure 3).
**Figure 3.** Schematic of the study design of Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT) intervention enrollment period, enrollment visit, postenrollment evaluation, and protocol refinement. The protocol is refined between each of the 3 waves of the RE-PACT study. JIT: just-in-time; mHealth: mobile health; REDCap: Research Electronic Data Capture; UCLA: University of California, Los Angeles; UW: University of Wisconsin–Madison.

**Total N=90**
- Total by site: American Family Children’s Hospital (45) UCLA (45)
- Total by arm: Intervention (n=45) Control (n=45)
  - Wave 1 (n=10) Wave 2 (n=20) Wave 3 (n=60)

**Enrollment Period**
(1-3 mo prior to T₀)
- Identify potential subjects via program rosters and chart review; mail families study invitation and information

**Waves 1-3**
(6 mo each)
**Enrollment Visit (T₀)**
- Obtain informed consent (in person at clinic visit or remotely via WebEx [Cisco] & REDCap [UW] survey); administer baseline study assessment

**Randomize**
- Intervention Arm
  - Respiratory illness action plan
  - Weekly mHealth SMS text messages
  - Monthly study assessments
  - Action planning and FIT coaching
  - Usual comprehensive medical care and coordination via complex care program

- Control Arm
  - Usual comprehensive medical care and coordination via complex care program

**Post-enrollment Evaluation**
(T₀ + 6 mo)
- Administer 6 mo study assessment

**Protocol refinement**
between each wave
- By each wave, refine RE-PACT protocol based on feasibility, acceptability, and fidelity measures:
  - Wave 1: focus on onboarding and training, recruitment, data collection
  - Wave 2: focus on randomization and intervention activities
  - Wave 3: focus on rapid enrollment and the conduct of all protocol activities with high-fidelity
Table 1. Schedule of activities of the Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT) intervention throughout the study period, with the depiction of personnel involved.

<table>
<thead>
<tr>
<th>Activities</th>
<th>Study period (personnel involved)</th>
<th>Enrollment visit (research coordinator)</th>
<th>RE-PACT intervention period (clinicians)</th>
<th>Final visit (research coordinator)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (T₀)ᵃ</td>
<td></td>
<td>1 (T₁)ᵃ</td>
<td>2 (T₂)ᵇ</td>
<td>3 (T₃)ᵇ</td>
</tr>
<tr>
<td>Confirm eligibility</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Informed consent</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Baseline assessment</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>6-mo assessment</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Randomization</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Participant compensation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Usual comprehensive medical care and coordination via complex care program</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓</td>
</tr>
</tbody>
</table>

**Intervention arm only**

- SMS text message training ✓
- Weekly mHealthᵇ text message and response ✓ ✓ ✓ ✓ ✓ ✓ ✓
- Intervention overview ✓
- Create action plan ✓
- Action planning ✓ ✓ ✓ ✓ ✓ ✓ ✓
- Rapid clinical response when triggered ✓ ✓ ✓ ✓ ✓ ✓ ✓
- Monthly study assessments ✓ ✓ ✓ ✓ ✓ ✓ ✓

ᵃMonth (time point).
ᵇmHealth: mobile health.

**Outcomes**

**Primary Study End Points: Feasibility, Acceptability, and Fidelity**

The specific measures and prespecified definitions of success for primary study end points, including feasibility, acceptability, and fidelity, are listed in Table 2. These measures will be summarized between each of the 3 waves, with protocol adjustments made for any measures that do not meet the definition of success.
Table 2. Primary study end points to evaluate the feasibility, acceptability, and fidelity of Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT; n=90).

<table>
<thead>
<tr>
<th>Measure and measure detail</th>
<th>Success definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feasibility</strong></td>
<td></td>
</tr>
<tr>
<td>Recruitment</td>
<td>Days to enroll target, mean (SD)</td>
</tr>
<tr>
<td>Intervention onset</td>
<td>Days between randomization and “T₀a” intervention activities, mean (SD)</td>
</tr>
<tr>
<td>Time to action plan</td>
<td>Days to action plan creation</td>
</tr>
<tr>
<td>Intervention time</td>
<td>Time logged (min) for action planning and for coaching activities, mean (SD)</td>
</tr>
<tr>
<td>Intervention costs</td>
<td>Mileage and travel costs; personnel salary; training costs; and other incurred costs, total</td>
</tr>
<tr>
<td>Intervention triggers</td>
<td>Number per patient (annualized), both respiratory and nonrespiratory focused</td>
</tr>
<tr>
<td><strong>Acceptability</strong></td>
<td></td>
</tr>
<tr>
<td>Enrollment</td>
<td>Enrollment rate (number of patients enrolled/number approached)</td>
</tr>
<tr>
<td>Consent refusal</td>
<td>Categorized reasons for refusal</td>
</tr>
<tr>
<td>Loss and dropouts</td>
<td>Dropout rate (active or passive) before 6 mo (number of dropouts/number enrolled)</td>
</tr>
<tr>
<td>Action plan, SMS text messaging and clinical responder satisfaction</td>
<td>Do caregivers use the action plan, coaching, and texting? How could it be improved? Would caregivers recommend this to another family?</td>
</tr>
<tr>
<td><strong>Fidelity</strong></td>
<td></td>
</tr>
<tr>
<td>Enrollment duration</td>
<td>Time (mo) of participant enrollment in the study, mean (SD)</td>
</tr>
<tr>
<td>Action plan creation</td>
<td>Number of respiratory and overall action plans per patient and action plan focus areas</td>
</tr>
<tr>
<td>Rapid clinical response: home or web-based visit</td>
<td>Success rate (number of visits completed and number expected); stratify by trigger and by “respiratory” and “nonrespiratory”</td>
</tr>
<tr>
<td>mHealth text messaging and Clinical responder satisfaction</td>
<td>Response rates (number of SMS texts responded and number expected); “respiratory” and “nonrespiratory”</td>
</tr>
<tr>
<td>Crossover</td>
<td>Number of patients inappropriately receiving any intervention component</td>
</tr>
<tr>
<td>Data collection</td>
<td>Complete entry and exit questionnaire, monthly questionnaire, and chart review data (number of data collection events completed and number of total data collection events)</td>
</tr>
</tbody>
</table>

^a: Enrollment visit  
^bN/A: not applicable.  
^c mHealth: mobile health.

**Clinical End Points**

The study’s secondary objective is to estimate the effect size of RE-PACT. The clinical end points are listed in Textbox 1. The primary clinical end point is severe respiratory illness, defined as a respiratory diagnosis requiring hospitalization. Respiratory diagnosis is defined as a discharge diagnosis of any of the following: asthma, pneumonia (community or hospital acquired), bronchiolitis, influenza, upper or lower respiratory tract infection, tracheitis, aspiration pneumonia and pneumonitis, chronic lung disease, and respiratory failure [34]. Hospitalization is defined as a nonelective, unscheduled hospital encounter (inpatient or observation status), accompanied by both an admission history and physical examination as well as a discharge summary note signed by a physician or advanced practice provider. Field-testing the assessment of this end point with trained research personnel at study sites demonstrated high interrater reliability (κ>0.9).
### Textbox 1. Clinical end points of Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT).

#### Primary clinical end point
- Severe respiratory illness, defined as a respiratory diagnosis requiring hospitalization

#### Secondary clinical end points
- Hospital days during severe respiratory illness
- Systemic steroid courses (systemic steroids [excluding inhaled or topical steroids for the purposes of defining an illness]: hydrocortisone, prednisone, prednisolone, dexamethasone, methylprednisolone, and triamcinolone acetonide; common inhaled steroids: fluticasone, budesonide, mometasone, beclomethasone, and triamcinolone [14-16,35])
- Systemic antibiotic courses (antibiotics: amoxicillin or amoxicillin/and clavulanate, ampicillin, ampicillin and sulbactam, azithromycin, cefdinir, cefepime, cefixime, cefpodoxime, ceftriaxone or cefotaxime, cefibuten, cefuroxime, cephalaxin [Keflex], clarithromycin, clindamycin, ciprofloxacin, doxycycline, erythromycin, erapenem, imipenem, levofloxacin, linezolid, meropenem, metronidazole, moxifloxacin, oseltamivir, penicillin, pipercillin and tazobactam, rifampin, and vancomycin [14-16,35])
- Respiratory emergency department visits
- Death

The secondary clinical outcomes (Textbox 1) include total hospital days during severe respiratory illness; the number of systemic steroid courses, systemic antibiotic courses, and respiratory ED visits; and death. Hospital days are calculated through resolution if admission occurs in the study time frame, even if discharge occurs after the study exit date. *Systemic corticosteroid course* is defined by oral or parenteral corticosteroids prescribed for respiratory diagnosis, including hydrocortisone, prednisone, prednisolone, methylprednisolone at least 1 mg/kg/d (or 30 mg/d) × minimum 3 days, or dexamethasone at least 0.15 mg/kg/d (or 10 mg/d) × ≥1 days. Physiologic or stress replacement doses in adrenal insufficiency are excluded. *Systemic antibiotic course* is defined by oral or parenteral antibiotics prescribed for respiratory diagnosis for a minimum of 3 days. The specific antibiotics are derived from the Infectious Diseases Society of America pediatric pneumonia guidelines [36] and published literature [35]. Respiratory ED visits are any ED visits not resulting in admission and having a discharge respiratory diagnosis.

### Exploratory Study End Points

The objectives of the tertiary study are to explore the mediating relationships between RE-PACT and capability, opportunity, motivation, and behavior (COM-B) measures [37]. By blending our foundational research on preventing hospitalizations [17,21] with behavioral intervention theory [38], our conceptual model suggests that decisions to seek care (behaviors) are influenced by capability (family capacity), opportunity (health system and susceptibility), and motivation (confidence). A theorized mechanism of RE-PACT’s effect is that combining action planning, mHealth surveillance, and coaching will increase caregiver COM-B measures to manage respiratory illness in severe CP. The tertiary end points are listed in Textbox 2.

### Textbox 2. Exploratory study end points of Respiratory Exacerbation–Plans for Action and Care Transitions (RE-PACT).

#### Capability
- Family Caregiver Activation in Transition (FCAT) tool [39]: mean composite score
- Caregiver General Self-Efficacy Scale (GSES) [40]: mean composite score

#### Opportunity
- Family Experiences with Care Coordination (FECC) [41]: percentage top-box score for selected items

#### Motivation
- Confidence responses mobile health SMS text messaging: mean weekly score

### Assessment Procedures

Data about research participants (children and their families) will be collected by study research assistants on case report forms using electronic family self-administered questionnaires, structured interviews with research personnel by telephone or in person with enrolled caregivers, and abstraction of child medical record data. Family and child measures will be recorded at baseline, and end points will be recorded at study exit (6 months after the enrollment visit [T0]). Caregiving measures, which may change as a result of the intervention, will be collected at baseline and study exit. Clinical responders will enter data for intervention group families into a clinical response event case report form.

Feasibility, acceptability, and fidelity end point data will be collected during each of the 3 waves by research personnel reviewing study logs, conducting monthly chart reviews, and administering surveys (by telephone, in person, or sending electronic self-administered links) with caregivers randomized to the intervention. For control group participants, the feasibility
of assessments will be evaluated by completion rates at study exit. In addition, intervention and control participants will be debriefed at study exit on their experiences in the study and asked for feedback on the strengths and weaknesses, as well as any concerns about the protocol. Between each wave and after the third wave, clinical teams at each site will be debriefed on the strengths and weaknesses, as well as concerns about the protocol.

The CP GMFCS measures and all caregiving measures have been well documented as reliable in the literature [39-45]. We have separately established the reliability of identifying respiratory illnesses in our preliminary research ($\kappa>0.9$). We will ensure reliability in data collection through direct observation, data auditing, establishing clear data dictionaries and definitions, using uniform variable definitions, and use of a central data repository coordinated and maintained by UW.

### Data Collection, Storage, and Protection

Clinical data (including adverse effects) and clinical laboratory data will be entered into Research Electronic Data Capture (REDCap; UW) managed by the University of Wisconsin, a 21 Code of Federal Regulation Part 11–compliant data capture system provided by the UW Institute for Clinical and Translational Research. The data system includes password protection and internal quality checks, such as automatic range checks, to identify inconsistent, incomplete, or inaccurate data. Clinical data will be entered directly from the source documents or entered directly from secure self-administered questionnaires (surveys) sent via REDCap to participants.

#### Sample Size Considerations

We will enroll up to 90 participants. On the basis of this team’s preliminary work, we estimate that half (45/90, 50%) of the participants will experience at least 1 respiratory illness during the enrollment period. We expect to be able to maintain contact and collect data from ≥90% (≥81/90) of the participants at the final follow-up, evenly divided between the intervention and control groups. We assume that this sample will not be powered to establish the efficacy of the intervention; however, it will provide a sufficient sample to determine feasibility and estimate effect sizes, which will be used for power calculations in the future large randomized clinical trial. Attainable power levels were calculated for detecting differences in severe respiratory illness rates (primary clinical outcome) between the study arms at the 2-tailed $<.05$ significance level based on a negative binomial (NB) regression model with an overdispersion parameter of $\phi=1.0$ (Table 3). Hence, large effect sizes with relative risks ranging between 3.0 and 5.0 for comparing the severe respiratory illness rates between the study arms will be detected with 19% to 88% power at the 2-tailed $<.05$ significance level.

#### Statistical Analysis Plan

The primary outcome data will assess the RE-PACT intervention’s feasibility, acceptability, and fidelity using descriptive statistics. Categorical variables will be displayed as percentages and continuous variables as means with SDs (if normally distributed) or medians with IQRs (if skewed). We will compare observed values with the prespecified definitions of success for each of our feasibility, acceptability, and fidelity measures. We will also determine overall positive, neutral, and negative reports of feasibility and acceptability using content analysis of qualitative (open-ended comments) data. We will explore any patterns if challenges emerge (eg, enrollment refusal or dropout or low reported use of the intervention activities).

We anticipate being underpowered to assess efficacy using the clinical end points of the intervention in this pilot study. Analyses will also estimate the effect size estimates of clinical end points to allow precise sample size calculations for a future large-scale efficacy trial. We will compare differences between the intervention and active control group outcomes at 6 months. The primary clinical outcome is the severe respiratory illness rate, defined as the total number of severe respiratory illnesses divided by the person-months over the 6-month follow-up period. The severe respiratory illness rate will be analyzed using an NB regression model to account for overdispersion in the count data. For the primary analysis, univariate NB regression analysis will be conducted with a study arm as a predictor variable. The study site will be included as a stratification factor in the primary analysis to account for stratified randomization. The observed effect size of the analysis will be quantified in terms of relative risk and reported along with the corresponding 95% CI. As a secondary analysis, multivariate NB regression analysis will be performed to compare the severe respiratory illness rates between the study arms. This analysis will include clinical and demographic characteristics as covariates in an initial nonparsimonious model. The least absolute shrinkage and selection operator and elastic net penalty methods for NB regression models will be used to identify a parsimonious model with independent covariates.

Longitudinal changes in the severe respiratory illnesses within and between study arms will be evaluated with a generalized linear mixed effects model with a logit link function and
participant-specific random effects. An autoregressive correlation structure will be used to account for within-participant correlations. In this analysis, the presence or absence of severe respiratory illness at the monthly assessments will be the dependent variable, the study arm will be included as a predictor variable, and the study site will be included as a stratification variable to account for the stratified randomization.

The secondary clinical outcomes include total hospital days during severe respiratory illness; the number of systemic steroid courses, systemic antibiotic courses, and respiratory ED visits; and death. The number of systemic steroid courses, systemic antibiotic courses, and respiratory ED visits over the 6-month follow-up period will be analyzed using NB regression analysis as described previously for the primary outcome. Observed effect sizes and the corresponding 95% CIs will be reported. The presence or absence of systemic steroid courses, systemic antibiotic courses, and respiratory ED visits will be documented at the monthly assessments, and longitudinal changes within and between study arms will be analyzed using generalized linear mixed effects modeling with a logit link function and patient-specific random effects. The total number of hospital days over the 6-month follow-up period will be analyzed using ANOVA with the study site as a stratification factor. In a secondary analysis, an analysis of covariance (ANCOVA) will be performed where clinical and demographic baseline characteristics will be included as covariates, and the least absolute shrinkage and selection operator method will be used to identify a parsimonious model. Longitudinal changes in the number of hospital days per hospitalization will be analyzed using a normal mixture linear mixed effects model with patient-specific random effects. The normal mixture component will be included in the model to capture the probabilities of hospitalization at the monthly follow-up. Parameter estimation will be performed using the expectation-maximization algorithm, the standard method for the parameter estimation of mixture models.

Two-tailed \( P \) values of <.05 will be considered statistically significant. Missing values (eg, owing to loss of follow-up and missing monthly visits) will be evaluated by conducting a sensitivity analysis comparing the results obtained from the complete case analysis with those obtained from imputation-based analyses. Specifically, multiple imputations will be used to impute the missing values of the primary and secondary clinical outcomes. For monotonic missing value data structures, we will use regression-based multiple imputation techniques. By contrast, we will use Markov Chain Monte Carlo–based imputation techniques for nonmonotonic missing value data structure.

Although we anticipate that the intervention and control groups will be similar owing to random assignment, we will adjust for any variables in our analysis that are not equal between the groups, given the small sample size. In addition, we will analyze for any effect of primary home language on the study outcomes because this may affect families’ ability to navigate the systems of care in the United States.

Finally, as a planned exploratory analysis, we will test the mediating effect of caregiver COM-B measures on the relationship between intervention and respiratory illness outcomes. The mediating effects will be evaluated by conducting a multistep analysis approach. In the initial step, NB regression analyses will be conducted to examine whether there are differences in respiratory illness outcomes (the number of severe respiratory illnesses, systemic steroid courses, systemic antibiotic courses, and respiratory ED visits) between the intervention and control arms. In the next step, we will conduct a sequence of univariate analyses by regressing each potential mediator variable (caregiver capability COM-B measures) on the binary study arm variable. If significant associations between the potential mediator variables and the study arm are detected, we will regress the respiratory illness outcomes on both the mediator variables and study arm indicator variables using ANCOVA. The mediation effect for each potential mediator variable will then be tested using the Sobel \( z \) test based on the slope parameter estimates from the corresponding regression models.

**Ethical Considerations**

This study received initial approval from the UW health system’s institutional review board on January 19, 2022 (20211532). All participants will provide informed consent before taking part in the study. Informed consent materials will be provided in private spaces in both written and verbal formats and will review in detail the study design, including random assignment to the intervention and control groups, potential risks of participation, protections against risk, and the rights of human research subjects. The informed consent process will also include review and signing of the Health Insurance Portability and Accountability Act waiver, allowing researchers to review the child’s medical records. Parents will be able to decline parts of the study and still participate in other parts and can revoke their consent at any point. Any identifying information kept for the purpose of contacting participants will be kept secure, in REDCap, a locked filing cabinet or in a password-protected electronic file and will be destroyed when the study is complete. The study is monitored by the Data Monitoring Committee at the UW-Madison Institute for Clinical and Translational Research. All participants receive an incentive of US $200, divided in 2 parts: US $100 at enrollment and US $100 after the exit survey, in the form of a gift card, check, or cash.

**Results**

The recruitment of the first wave began on April 27, 2022. To date, we have enrolled 30 (33%) out of 90 participants, as projected. The final wave of recruitment will end by October 31, 2023, and the final participant will complete the study by April 30, 2024. We will start analyzing the complete responses by April 30, 2024, and the publication of results is expected at the end of 2024.

**Discussion**

**Summary**

We describe the protocol for a pilot clinical trial of RE-PACT, a JIT adaptive intervention to reduce respiratory illness in severe
CP. A recent expert consensus statement on preventing and managing respiratory disease in young people with CP highlighted the need for 4 activities: early identification of risk factors; regular assessment of risk; effective partnerships among multidisciplinary teams, families, and individuals with CP; and proactive treatment of respiratory disease [46]. The RE-PACT intervention protocol aligns with each of these 4 critical areas.

For children with severe CP, RE-PACT was designed by families and clinicians from promising earlier interventions to manage health crises with proactive action planning, simple surveillance of family confidence to avoid hospitalization through frequent SMS text messaging, and JIT adaptive rapid clinical responses. This intervention breaks down barriers to equitably connect families and clinical teams precisely when it matters most. This approach is innovative because we tailor the intensity of the response (eg, telephone call and clinic visit) and its content to family- and illness-specific needs. The adaptive nature of the intervention ensures that it meets caregiver needs for that specific instance, flexibly changing for individuals over time in response to each intervention trigger. RE-PACT is also designed to acknowledge that respiratory illness in severe CP is driven by both respiratory and nonrespiratory comorbid and social conditions [11,46] (eg, neuromuscular weakness, seizures, dysphagia, feeding intolerance, health system navigation barriers, and coordination problems).

By conducting successively larger waves of the RE-PACT protocol, we expect to produce a final high-quality protocol that has been developed sufficiently to support the implementation of a large-scale multisite clinical efficacy trial.

Limitations
This study has several limitations. Although we anticipate achieving a feasible, acceptable, and high-fidelity protocol by the end of the third wave, it is possible that some challenges may remain. We anticipate being underpowered to assess intervention efficacy. Despite the randomized design, allocation concealment is not possible. As randomization occurs at the level of the family, inadvertent intervention contamination to nonintervention patients in the same clinical program may occur. This risk will be minimized by having research staff (not clinical staff) manage action planning and SMS text messaging procedures. In future studies, we will consider alternative designs (such as a stepped wedge trial), randomizing at the clinic level to avoid this threat. Threats to external validity will reflect the relatively narrow population of families recruited from 2 complex care programs. Although it is a strength that the intervention will be conducted in English and Spanish, future expansion to populations of children with CP outside of complex care programs and from more geographically and culturally diverse settings will be helpful. As research continues, it will be important to examine whether this intervention design, which relies in part on the use of digital technology, addresses disparities in access to care and inequities in outcomes.

Conclusions
Despite the limitations, our pilot RE-PACT intervention represents an innovative and promising strategy to reduce severe respiratory illness among children with severe CP. RE-PACT operationalizes universal action planning, mobile SMS text messaging, and a JIT adaptive rapid clinical response to deliver timely customized care to families of children with severe CP. This protocol describes detailed methods to assess intervention feasibility, acceptability, and fidelity. This line of research may be relevant to other researchers and health care providers who wish to adopt a similar early intervention strategy for patients with chronic and complex conditions at high risk of future hospitalization.

Acknowledgments
The research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (R34HL153570). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability
The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors’ Contributions
RC, the principal investigator of this study, obtained grant funding and conceived the study. CL, HK, TK, PC, CC, DG, BK, ME, and SI participated in the design of the study. AF and RC drafted the manuscript. RC, CL, GW, KH, LP, RD, SI, and TW are responsible for recruitment and major study activities. All authors contributed to the intellectual content of the manuscript and the development of the trial protocol, and all authors have read, revised, and approved the final manuscript.

Conflicts of Interest
None declared.

References

https://www.researchprotocols.org/2024/1/e49705

JMIR Res Protoc 2024 | vol. 13 | e49705 | p.141
(page number not for citation purposes)


Abbreviations

ACTIV: Assessing Confidence at Times of Increased Vulnerability
ANCOVA: analysis of covariance
COM-B: capability, opportunity, motivation, and behavior
CP: cerebral palsy
ED: emergency department
GMFCS: Gross Motor Function Classification System
ICD-10: International Classification of Diseases, Tenth Revision
JIT: just-in-time
mHealth: mobile health
NB: negative binomial
PACT: Plans for Action and Care Transitions
RE-PACT: Respiratory Exacerbation–Plans for Action and Care Transitions
REDCap: Research Electronic Data Capture
UCLA: University of California, Los Angeles
UW: University of Wisconsin–Madison

©Alyssa Fleischman, Carlos Lerner, Heidi Kloster, Paul Chung, Thomas Klitzner, Christopher Cushing, Danielle Gerber, Barbara Katz, Gemma Warner, Kristina Devi Singh-Verdeflor, Roxana Delgado-Martinez, Lorena Porras-Javier, Siem Ia, Teresa Wagner, Mary Ehlenbach, Ryan Coller. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.01.2024. 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.
Efficacy and Safety of the Natural Killer T Cell–Stimulatory Glycolipid OCH-NCNP1 for Patients With Relapsing Multiple Sclerosis: Protocol for a Randomized Placebo-Controlled Clinical Trial

Tomoko Okamoto, MD, PhD; Takami Ishizuka, PhD; Reiko Shimizu, MD, PhD; Yasuko Asahina, PhD; Harumasa Nakamura, MD, PhD; Yuko Shimizu, MD, PhD; Yoichiro Nishida, MD, PhD; Takanori Yokota, MD, PhD; Youwei Lin, MD; Wakiro Sato, MD, PhD; Takashi Yamamura, MD, PhD

1Department of Neurology, National Center Hospital, National Center of Neurology and Psychiatry, Tokyo, Japan
2Department of Clinical Research Support, Clinical Research and Education Promotion Division, National Center Hospital, National Center of Neurology and Psychiatry, Tokyo, Japan
3Department of Neurology, Tokyo Women's Medical University, Tokyo, Japan
4Department of Neurology and Neurological Science, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan
5Department of Immunology, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan

Corresponding Author:
Takashi Yamamura, MD, PhD
Department of Immunology
National Institute of Neuroscience
National Center of Neurology and Psychiatry
4-1-1 Ogawahigashi-cho, Kodaira
Tokyo, 187-8502
Japan
Phone: 81 42 341 2712 ext 5241
Email: yamamura@ncnp.go.jp

Abstract

Background: Multiple sclerosis (MS) is an autoimmune inflammatory disease of the central nervous system that causes myelin sheath damage and axonal degeneration. The glycolipid (2S, 3S, 4R)-1-O-(α-d-galactosyl)-2-tetracosanoylamino-1,3,4-nonaetriol (OCH-NCNP1 or OCH) exerts an immunoregulatory action that suppresses T helper (Th)1 cell–mediated immune responses through natural killer T cell activation, selective interleukin-4 production, and Th2 bias induction in human CD4-positive natural killer T cells.

Objective: This trial aims to investigate the efficacy and safety of the immunomodulator OCH in patients with relapsing MS through 24-week repeated administration.

Methods: This protocol describes a double-blind, multicenter, placebo-controlled, randomized phase II clinical trial that was initiated in September 2019. The participants were randomly assigned to either a placebo control group or an OCH-NCNP1 group and the investigational drug (3.0 mg) was orally administered once weekly for the 24-week duration. Major inclusion criteria are as follows: patients had been diagnosed with relapsing MS (relapsing-remitting and/or secondary progressive MS) based on the revised McDonald criteria or were diagnosed with MS by an attending physician as noted in their medical records; patients with at least two medically confirmed clinical exacerbations within 24 months prior to consent or one exacerbation within 12 months prior to consent; patients with at least one lesion suspected to be MS on screening magnetic resonance imaging (MRI); and patients with 7 points or less in the Expanded Disability Status Scale during screening. Major exclusion criteria are as follows: diagnosis of neuromyelitis optica and one of optic neuritis, acute myelitis, and satisfying at least two of the following three items: (1) spinal cord MRI lesion extending across at least three vertebral bodies, (2) no brain MRI lesions during onset (at least four cerebral white matter lesions or three lesions, one of which is around the lateral ventricle), and (3) neuromyelitis optica–immunoglobulin G or anti-aquaporin-4 antibody-positive. Outcome measures include the primary outcome of MRI changes (the percentage of subjects with new or newly expanded lesions at 24 weeks on T2-weighted MRI) and the secondary outcomes annual relapse rate.
Multiple sclerosis (MS) is considered to be an autoimmune disease triggered by environmental factors that act on a genetically susceptible host. Both major histocompatibility complex (MHC) and non-MHC genes are risk factors for the development of MS. In addition, environmental factors such as low vitamin D, low ultraviolet radiation exposure, cigarette smoking, obesity, and Epstein-Barr virus exposure can increase the risk for both the development of MS and a more severe disease course. Accumulating evidence suggests that dysregulation of the intestinal microbiome (dysbiosis) constitutes an important factor contributing to MS pathogenesis. The microbiome regulates T cell function, with both regulatory and pathogenic effects, thereby playing an important role in autoimmune responses [1].

MS is a cell-mediated autoimmune disease directed against central nervous system (CNS) myelin antigens involving both CD4+ and CD8+ T cells, especially the so-called pathogenic T helper (Th)17- and Th1-type and CD8 myelin autoreactive T cells. The development of MS is likely triggered or promoted by breakdown of the delicate balance between autoreactive T cells and regulatory lymphocytes [1].

Although MS has historically been considered a demyelinating disease of the CNS and white matter, in recent years, neurodegeneration of the cortical and deep gray matter has been recognized to play a role in the pathogenesis of MS. Cortical atrophy is associated with disease progression, which has emerged as one of the best predictors of neurological disability in MS [1].

According to a recent worldwide epidemiological study, the number of patients with MS is estimated to be 2.8 million, which has been increasing in every world region since 2013 [2]. The estimated total economic burden of MS in the United States is US $85.4 billion, with a direct medical cost of US $63.3 billion and indirect and nonmedical costs of US $22.1 billion [3].

Relapsing-remitting MS (RRMS) initially involves clinical relapses with near or complete recovery; however, recovery over time may be incomplete and disability often worsens [4]. Approximately 20% of patients with RRMS develop a progressive form of MS accompanied by chronic neuroinflammation. Such cases are referred to as secondary progressive MS (SPMS). The term “relapsing MS” (RMS) is used to describe the condition of patients with repeated relapses of either RRMS or SPMS.

As understanding of the pathomechanism of MS progresses, various disease-modifying drugs have been used in clinical practice, including sphingosine-1-phosphate receptor modulators (fingolimod, siponimod), a monoclonal antibody that selectively binds the α4 integrin subunit (natalizumab), and CD20 monoclonal antibodies (ofatumumab, ocrelizumab), resulting in an overall improvement of patient prognosis. However, there are patients for whom current treatments are ineffective and there are cases where current treatments are intolerable due to side effects. Furthermore, progressive MS remains refractory to current drugs and constitutes an unmet medical need [4]. These observations highlight the need for new, safer therapeutic oral agents.

In 1997, Kawano et al [5] reported the discovery of sponge-derived α-galactosylerceramide (α-GC) as a glycolipid ligand that stimulates natural killer T (NKT) cells. However, because this glycolipid stimulates NKT cells and induces the production of both interleukin (IL)-4 and interferon (IFN)-γ, a search for glycolipids that selectively induce IL-4 production was initiated to treat Th1 cytokine-dependent autoimmune diseases such as MS. These efforts led to the discovery of OCH-NCNP1, a monotaurine, nonanetriol (OCH-NCNP1; hereafter referred to as OCH), which was detected during screening for glycolipids that selectively induce the production of Th2 cytokines [6].

OCH is a derivative (synthetic glycolipid) of α-GC with a shortened fatty-acid chain (sphingosine chain). The compound is a white to slightly yellow powder that is extremely insoluble in methanol or ethanol and is practically insoluble in water. OCH exhibits an immunoregulatory action that suppresses Th1 cytokine–dependent autoimmune responses through NKT cell activation, including sphingosine-1-phosphate receptor modulators (fingolimod, siponimod). OCH is a derivative (synthetic glycolipid) of α-GC with a shortened fatty-acid chain (sphingosine chain). The compound is a white to slightly yellow powder that is extremely insoluble in methanol or ethanol and is practically insoluble in water. OCH exhibits an immunoregulatory action that suppresses Th1 cytokine–dependent autoimmune responses through NKT cell activation.
selective IL-4 production, and Th2 bias induction in human CD4+ NKT cells [7].

Although the mechanism by which OCH induces Th2 bias in NKT cells is not fully understood, it has been suggested to involve both cell intrinsic and extrinsic factors [8,9]. As a cell intrinsic mechanism, IFN-γ production by NKT cells was reported to be more susceptible to the sphingosine length of the glycolipid ligand compared to IL-4 production, and the length of the sphingosine chain determined the half-life of NKT cell stimulation by CD1d-associated glycolipids. As an extrinsic regulatory mechanism, OCH suppresses the production of IFN-γ, not only by NKT cells but also by NK cells, compared with that of α-GC. OCH induced lower IL-12 production due to ineffective primary IFN-γ and CD40 ligand expression by NKT cells, which resulted in lower secondary IFN-γ induction.

Animal studies verified that OCH can be administered orally to control Th1 cell–mediated autoimmune pathology in mice, with efficacy against autoimmune disease models such as experimental autoimmune encephalomyelitis, collagen-induced myelitis [10], nonobese diabetes [11], and inflammatory enteritis [12].

We here describe a protocol for a feasibility study that will be conducted in patients with RMS to investigate the efficacy and safety of OCH immunomodulators. Through 24 weeks of weekly administration, this trial was performed to confirm changes in exploratory T cell and NK cell biomarkers that fluctuated in the phase I trial [13], as well as measures of efficacy, including recurrence, dysfunction, and magnetic resonance imaging (MRI) changes, and their association with combined endpoints. In alignment with past clinical trials [14,15] and European Medical Agency guidelines, the primary outcome will be the percentage of subjects with new or newly expanded lesions at 24 weeks on T2-weighted MRIs.

**Methods**

**Study Design**

This is a double-blind, multicenter, placebo-controlled, randomized phase II clinical trial with a 24-week duration. This trial investigates the efficacy and safety of the immunomodulator OCH in patients with RMS through 24-week repeated administration. The protocol was also designed to confirm exploratory biomarkers of T cells and NK cells that fluctuated in the phase I trial [13], as well as efficacy-related clinical endpoints, including relapses, disability, and MRI changes, and their association with combined endpoints. Recruitment opened in September 2019 and concluded in June 2021, with participants randomly assigned to either a placebo control group or an OCH group using a clinical-based management system (Translational Research Center for Medical Innovation, Kobe, Japan). The researchers were blinded to the participant group assignments, and the packaging appearance of the control and investigational drugs was confirmed to be indistinguishable.

An independent data monitoring committee will regularly audit all available data, including safety, and recommend trial continuation to the principal investigator.

**Selection Criteria**

Inclusion and exclusion criteria for this study are provided in Textbox 1.
Textbox 1. Inclusion and exclusion criteria.

### Inclusion criteria

- Patients who provided written consent for trial participation
- Patients had been diagnosed with relapsed multiple sclerosis (MS), including relapse-remitting MS and/or secondary progressive MS, based on the revised McDonald criteria or were diagnosed with MS by an attending physician as noted in their medical records
- Patients with at least two medically confirmed clinical exacerbations within 24 months prior to consent or one exacerbation within 12 months prior to consent
- Patients with at least one lesion suspected to be MS on screening magnetic resonance imaging (MRI)
- Patients with 7 points or less in the Expanded Disability Status Scale during screening
- Patients who were 20-65 years old at the time of consent
- Patients who can practice contraception until 90 days after final administration of the investigational drug
- Patients with no clinical or test findings suggesting acute recurrence based on evaluation by a neurologist

### Exclusion criteria

- Diagnosis of neuromyelitis optica and one of the three following three criteria: optic neuritis, acute myelitis, and satisfied at least two of the following three items: (1) spinal cord MRI lesion extending across at least three vertebral bodies, (2) no brain MRI lesions during onset (at least four cerebral white matter lesions or three lesions, one of which is around the lateral ventricle), and (3) neuromyelitis optica–immunoglobulin G (IgG) or anti-aquaporin-4 antibody-positive
- Currently pregnant or nursing
- Contraindication for MRI (eg, those with metal implants or a pacemaker) or cases in which MRI is difficult to perform (eg, claustrophobia)
- History of allergic or hypersensitive reaction to gadolinium contrast
- History of liver disease, including liver transplantation, viral hepatitis, autoimmune hepatitis, cirrhosis, and hepatic malignancies
- Liver dysfunction determined by the screening test or baseline test (alanine aminotransferase, aspartate aminotransferase, γ-guanosine triphosphate, or alkaline phosphatase) exceeding twice the upper limit of normal and total bilirubin exceeding 1.5-fold the upper limit of normal
- History of malignant tumors in the past 5 years (however, patients deemed to have no recurrence for at least 5 years prior to consent can be enrolled)
- Varicella-zoster virus IgG antibody–negative
- Positive for syphilis serum reaction (treponema pallidum latex agglutination, rapid plasma reagin)
- Positive for β-D glucan (exceeding the standard value) or T-spot
- Positive for anti-aquaporin-4 antibody
- History of human immunodeficiency virus infection or who have been confirmed positive by a screening test
- History of hepatitis B infection (hepatitis B surface antigen–positive or hepatitis B core antibody–positive), or patients who have been confirmed positive by a screening test
- History of stem cell transplantation, organ transplantation, and treatment for rejection
- Physical, mental, or social condition affecting the ability to provide consent to or complete the trial
- Participation in other clinical trials that received the investigational drug within 4 months prior to enrollment
- Blood donation (200 mL within 2 months, 400 mL within 3 months) prior to enrollment
- Peripheral blood lymphocyte count <600/mm<sup>3</sup> by screening or baseline examination
- With or suspected of having an infectious disease
- Immunocompromised patients
- Inflammatory bowel disease
- Use of the following prohibited concomitant therapies:
Sample Size Estimate
The target number of participants was 30 (15 per group) based on feasibility. The detection power was calculated when the primary outcome (proportion of subjects with new or enlarged existing lesions detected on T2-weighted images) was compared using the Fisher exact test. With 60% of patients placed in the placebo group based on the results of the domestic phase II study (D1201 study) at the time of fingolimod development [16] and assuming a 5% proportion in the OCH group, the detection power was 86.5%. Furthermore, when the OCH group was set to 10%, the detection power was 74.6% (the significance level was set to 5% using a two-tailed test).

Patient Clinodemographic Characteristics
The following clinodemographic information was recorded for each participant: (1) subject background (eg, birth date, sex, body height, body weight); (2) urine pregnancy test; (3) infectious disease or antibody test; (4) vital signs (blood pressure, pulse rate, body temperature, and breathing rate); (5) neurological symptoms rating scale (Expanded Disability Status Scale [EDSS] or functional disability scale); (6) clinical tests (hematological test, hematobiochemical test, or urine test); (7) 12-lead electrocardiogram; (8) echocardiography; (9) chest or abdominal X-ray; (10) abdominal computed tomography; (11) MRI; (12) peripheral blood gene expression level (reverse transcription–polymerase chain reaction); (13) frequencies of NK, NKT, T cells, or other lymphocyte subsets; (14) frequencies of Th1, Th2, or Th17 cells; (15) intestinal and oral microbiome; (16) adverse events; (17) combination drugs and important nondrug therapies; and (18) Columbia Suicide Severity Rating Scale.

Outcome Measurements and Study Timeline
The outcomes that will be measured repeatedly throughout the trial and the corresponding time points are shown in Multimedia Appendix 1. The main outcome is MRI changes (the percentage of subjects with new or newly expanded lesions on T2-weighted MRI). In addition, the annual relapse rate, cumulative number of new or newly expanded lesions on T2-weighted MRI, percentage of subjects who did not have lesions at 12 and 24 weeks on T2-weighted images of head MRI compared to those at the preobservation period at 24 weeks, brain atrophy in T1-3D images, percentage of subjects with no contrast-enhanced lesions on gadolinium-enhanced T1-weighted images, number of contrast-enhanced lesions on gadolinium-enhanced T1-weighted images, changes in volume of demyelinating lesions on fluid-attenuated inversion recovery 3D images, diffusion tensor imaging changes in myelin sheath lesions, and changes in myelin sheath lesions by a myelin map will be assessed. Additional outcomes include the relapse-free period (from randomization to earliest relapse), sustained reduction in disability occurrence rate, period until sustained reduction in disability (from randomization), no evidence of disease activity, and exploratory biomarkers from phase I trials (peripheral blood gene expression, frequencies of lymphocyte subsets, and intestinal and oral microbiome) [13].

Adverse Events
Adverse events, defined as any undesired or unintended sign (including abnormal fluctuations in each test value), symptoms, or illness that occurs between the start of investigational drug administration and end or discontinuation of the investigational drug (regardless of its association with the study drug), will be recorded. The recordings will include the degree of symptoms on a 1-5 scale (1=asymptomatic or mild adverse events, 5=death due to the adverse event), outcomes from 1 to 6 (1=event disappeared or normalized, 6=death), and association with the
investigational drug from 1 to 4 (1=associated, 4=not associated).

**Discontinuation Criteria**

Discontinuation of involvement in the clinical trial may occur for any patient because of voluntary withdrawal or for any of the following reasons: (1) serious adverse events; (2) pregnancy; (3) alanine aminotransferase and aspartate aminotransferase >5-fold higher than the standard value upper limit; (4) total bilirubin exceeding 2.0 mg/dL; (5) number of peripheral blood lymphocytes <500/mm$^3$ and administration of the investigational drug is stopped three times in a row without recovery, even after three tests; (6) occurrence of a grade-3 event based on Common Terminology Criteria for Adverse Events v4.0-NCI; (7) new neurological symptoms with unpredictable MRI findings observed from the course of MS; (8) more than one recurrence of MS; (9) at least three courses of pulse steroid therapy performed; (10) an important management problem is discovered (subject noncompliance); (11) significant deviation from the protocol; (12) administration of the investigational drug is stopped at least three times, regardless of the reason; and (13) at the discretion of the principal investigator. If discontinuation occurs because of this final criterion, justification will be provided in the results report.

**Administration of Trial Compound**

The investigational drug OCH (3.0 mg) in the form of granules (total of 0.3 g) or placebo granules alone (0.3 g) was orally administered with approximately 150 mL of water 30 min before breakfast once weekly for the 24-week trial duration. The white granules were composed of crystalline cellulose, mannitol, sodium croscarmellose, low-substituted hydroxypropyl cellulose, and polysorbate 80.

**Statistical Methods**

Patient clinicopathological data will be summarized using descriptive statistics for each group. The proportion of subjects with new or enlarged existing lesions on T2-weighted MRIs in each group will be calculated and compared using the Fisher exact test with 95% CIs. The proportion of data between groups related to the primary endpoint will also be compared using the Fisher exact test with 95% CIs. For time-to-event data such as relapse-free periods, Kaplan-Meier estimates will be calculated for each group and groups will be compared using the log-rank test. The expression rates of safety-related outcomes and adverse events will be calculated for each group. For various test values, a list of measured values will be created for each group. The summary statistics for each group and measurement time point will also be calculated.

**Patient and Public Involvement**

Patients were first involved in this study during the recruitment and screening processes. Based on previous experiences with patients with MS, the importance of the outcomes measured in this study was evaluated and determined based on their medical importance and impact on the patients’ quality of life. The patients recruited for the study agreed to the methods of disseminating the aggregate results when providing informed consent.

---

**Results**

The first patient completed registration in December 2019 and the last patient completed registration in June 2021. The full analysis set comprised 30 cases and the study analysis was completed in March 2023. Preliminary analysis suggests that OCH may be effective for RMS (particularly SPMS).

**Discussion**

**Principal Results**

This will be the first randomized double-blind placebo-controlled trial to study the efficacy and safety of OCH. This randomized controlled trial will determine whether OCH is effective and safe in patients with MS. The results of this trial are expected to provide evidence for the potential of using OCH as a therapeutic agent for MS.

A single-dose trial (STEP 1) trial was conducted in healthy adults from November 2012 to July 2013 and a repeated-dose trial (STEP 2) was conducted in patients with RRMS from March 2014 to August 2017 [13]. In a phase I trial consisting of STEP 1 and STEP 2, OCH was administered to 28 patients (STEP 1: 3 patients×5 groups in all cohorts, STEP 2: 7 patients×1 group in a cohort, 3 patients×2 groups in a cohort).

Grade-1 adverse events and side effects were noted in STEP 1 and there were no serious adverse events or discontinuations. In STEP 2, serious side effects (acidosis and altered state of consciousness, depression, muscle weakness, and malaise) were reported in three patients in the 0.3 and 3 mg groups. Two patients discontinued treatment; however, all patients recovered. No other serious events were observed, confirming that the patients tolerated doses of 0.3, 1, and 3 mg. After OCH administration, there was no significant change in the neurological symptom evaluation scale score (EDSS or functional disability scale) even in STEP 2 because of the short observation period. MRI revealed clinically significant abnormalities in one patient in the 0.3 mg group; this patient discontinued the study.

A phase I physician-led clinical trial, conducted as an early exploratory clinical trial, determined that the dose was likely to have pharmacological action (fluctuations in some biomarkers) in humans. The dose at which pharmacological action was observed in experimental autoimmune...
encephalomyelitis model mice was 0.4-0.5 mg/kg. Considering that the area under the curve (AUC) following oral administration of 5 mg/kg to mice was 2922 ng·h/mL, the AUC following oral administration of 0.5 mg/kg (dose at which pharmacological action was observed) was estimated as 292.2 ng·h/mL. Assuming a correlation between systemic exposure (AUC) and pharmacological effects and systemic exposure in monkeys (AUC after oral administration of 10 mg/kg was 2376, SD 1164 ng), pharmacological action may be expected in humans if at least approximately 1.2 mg/kg is orally administered.

In STEP 1 of the phase I trial, administration of OCH at doses of 0.3, 1, 3, 10, 30, 100, and 300 mg to healthy adults was planned. However, STEP 1 was completed at 30 mg; in the subsequent STEP 2, OCH at 0.3, 1, and 3 mg was administered to patients with MS. Importantly, no tolerability problems were noted.

The following changes in exploratory biomarkers were observed. In analysis of the immune cell subsets using flow cytometry, (1) inhibitory T cells (Foxp3+T cells) and effector regulatory T cells (CD45RA-FoxP3\textsuperscript{high} T cells) tended to increase and (2) granulocyte-macrophage colony-stimulating factor–producing Th cells transiently decreased in both healthy subjects and patients with MS. Recently, granulocyte-macrophage colony-stimulating factor–producing Th cells were identified as pathogenic cells in MS [17]. These changes suggest that oral OCH administration can correct the proinflammatory changes linked with disease activity in MS. Moreover, by conducting DNA microarray analysis of whole blood cells, we identified upregulation of several immunoregulatory genes and downregulation of several inflammatory genes in both healthy subjects and patients with MS, further supporting the immunoregulatory effect of OCH.

Based on the above safety profile and biomarker analysis results, a dose of 3 mg was selected as the investigational dose for this trial.

**Limitations**
The primary limitation of our trial design is the small sample size. The sample size in this study was determined to be 30 participants based on the results of the previous phase I study. This study was limited to a small cohort of patients over a 24-week timeline and involved weekly administration of one dose of OCH. Although necessary in this early stage of the investigation, the small sample size could limit the ability to identify potential adverse events that may be rare or related to specific clinico-demographic traits of patients not captured in this study. However, the robust and clinically relevant nature of our primary outcome measure and sample size determined in prior studies are expected to provide indications of drug efficacy.

**Conclusions**
This article describes an NKT cell–stimulatory glycolipid (OCH) protocol. This randomized controlled trial will determine whether OCH is effective and safe in patients with MS. The results of the trial are expected to provide evidence for the potential of OCH as a therapeutic agent for MS.
References


Abbreviations

α-GC: α-galactosylceramide
AUC: area under the curve
CNS: central nervous system
EDSS: expanded disability status scale
IFN: interferon
IL: interleukin
MHC: major histocompatibility complex
MRI: magnetic resonance imaging
MS: multiple sclerosis
NKT: natural killer T cells

https://www.researchprotocols.org/2024/1/e46709

JMIR Res Protoc 2024 | vol. 13 | e46709 | p.152
(page number not for citation purposes)
OCH: (2S, 3S, 4R)-1-O-(α-d-galactosyl)-2-tetracosanoylamino-1,3,4-nonaetriol
RMS: relapsing multiple sclerosis
RRMS: relapsing-remitting multiple sclerosis
SPMS: secondary progressive multiple sclerosis
Th: T helper
Transcranial Magnetic Stimulation of the Default Mode Network to Improve Sleep in Individuals With Insomnia Symptoms: Protocol for a Double-Blind Randomized Controlled Trial

Lindsey Hildebrand1*, MA; Alisa Huskey1*, PhD; Natalie Dailey1, PhD; Samantha Jankowski1, BS; Kymberly Henderson-Arredondo1, BS; Christopher Trapani2, BS; Salma Imran Patel3, MD; Allison Yu-Chin Chen2, MD, PhD; Ying-Hui Chou2, ScD; William D S Killgore1, PhD

1Department of Psychiatry, University of Arizona, Tucson, AZ, United States
2Department of Psychology, University of Arizona, Tucson, AZ, United States
3Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, University of Arizona, Tucson, AZ, United States

*these authors contributed equally

Corresponding Author:
Lindsey Hildebrand, MA
Department of Psychiatry
University of Arizona
1501 N Campbell Ave
Tucson, AZ, 85724
United States
Phone: 1 651 410 9663
Email: hildebrandll@arizona.edu

Abstract

Background: Cortical hyperarousal and ruminative thinking are common aspects of insomnia that have been linked with greater connectivity in the default mode network (DMN). Therefore, disrupting network activity within the DMN may reduce cortical and cognitive hyperarousal and facilitate better sleep.

Objective: This trial aims to establish a novel, noninvasive method for treating insomnia through disruption of the DMN with repetitive transcranial magnetic stimulation, specifically with continuous theta burst stimulation (cTBS). This double-blind, pilot randomized controlled trial will assess the efficacy of repetitive transcranial magnetic stimulation as a novel, nonpharmacological approach to improve sleep through disruption of the DMN prior to sleep onset for individuals with insomnia. Primary outcome measures will include assessing changes in DMN functional connectivity before and after stimulation.

Methods: A total of 20 participants between the ages of 18 to 50 years with reported sleep disturbances will be recruited as a part of the study. Participants will then conduct an in-person screening and follow-on enrollment visit. Eligible participants then conduct at-home actigraphic collection until their first in-residence overnight study visit. In a double-blind, counterbalanced, crossover study design, participants will receive a 40-second stimulation to the left inferior parietal lobule of the DMN during 2 separate overnight in-residence visits. Participants are randomized to the order in which they receive the active stimulation and sham stimulation. Study participants will undergo a prestimulation functional magnetic resonance imaging scan and a poststimulation functional magnetic resonance imaging scan prior to sleep for each overnight study visit. Sleep outcomes will be measured using clinical polysomnography. After their first in-residence study visit, participants conduct another at-home actigraphic collection before returning for their second in-residence overnight study visit.

Results: Our study was funded in September 2020 by the Department of Defense (W81XWH2010173). We completed the enrollment of our target study population in the October 2022 and are currently working on neuroimaging processing and analysis. We aim to publish the results of our study by 2024. Primary neuroimaging outcome measures will be tested using independent components analysis, seed-to-voxel analyses, and region of interest to region of interest analyses. A repeated measures analysis of covariance (ANCOVA) will be used to assess the effects of active and sham stimulation on sleep variables. Additionally, we will correlate changes in functional connectivity to polysomnography-graded sleep.
Conclusions: The presently proposed cTBS protocol is aimed at establishing the initial research outcomes of the effects of a single burst of cTBS on disrupting the network connectivity of the DMN to improve sleep. If effective, future work could determine the most effective stimulation sites and administration schedules to optimize this potential intervention for sleep problems.

Trial Registration: ClinicalTrials.gov NCT04953559; https://clinicaltrials.gov/ct2/show/NCT04953559

International Registered Report Identifier (IRRID): DERR1-10.2196/51212

KEYWORDS
continuous theta burst stimulation; transcranial magnetic stimulation; default mode network; sleep; insomnia; cTBS; randomized controlled trial

Introduction

Overview

The sleep disturbances of insomnia include difficulty falling asleep, staying asleep, or inability to fall back to sleep after premature awakenings, or any combination of these, along with daytime sleepiness and dysfunction. Difficulty falling and remaining asleep affected nearly 30% of the US population, with approximately 20% of the general population experiencing occasional difficulties and an additional 10% meeting the clinical criteria for an insomnia disorder [1]. However, this protocol was conducted during the COVID pandemic, when the prevalence of insomnia symptoms increased to 53% and insomnia disorder prevalence jumped from 10% to 17% [2]. Insomnia symptoms have been linked with poorer health across multiple domains and exacts a high societal cost in the health care system, on the workplace, on the health care system, and on an individual’s and family’s quality of life [3-5]. Moreover, side effects from medications used to manage insomnia increase these factors in some cases. Novel approaches that facilitate sleep are critically needed, and some potential neural mechanisms of insomnia create a viable treatment target [6].

The most widely accepted theory of primary insomnia is the hyperarousal theory, which suggests that problems with sleep initiation and maintenance are due largely to the disruptive effects of somatic or cognitive hyperarousal [7,8]. In particular, cortical hyperarousal can emerge from excessive focus on repetitive negative thoughts, including intensive problem-solving, self-reflective rumination, and worry [9-11]. Subjectively, people who struggle with insomnia often make comments such as “I wish I could just turn my mind off” or “I just keep replaying conversations over in my head.” In fact, one of the major features of insomnia is the tendency toward self-reflective rumination and worry [9,12-15]. This internal dialogue contributes to a cycle of self-referential thought and hyperarousal that appears to hinder sleep onset and maintenance [9,16-18]. In fact, worry and sleep disruptions are associated in the general population [3]. In individuals with insomnia, cortical hyperarousal contributes to difficult initiating and maintaining sleep [19].

Neuroimaging research has shown that internally focused self-reflective processes of this type tend to activate an interconnected system within the brain known as the default mode network (DMN) [20,21]. Moreover, negative ruminative thinking is associated with changes in DMN connectivity and other brain regions associated with cognitive arousal or negative emotion [22]. Patients with insomnia often show abnormalities in the functioning of the DMN that are consistent with the hyperarousal hypothesis [23-25]. The core nodes of the DMN typically include the medial prefrontal cortex, posterior cingulate cortex, precuneus, and bilateral inferior parietal cortex regions but can also include several ancillary smaller regions, including the hippocampus, medial temporal lobes, and other subcortical structures [26]. Individuals with insomnia disorder tend to have increased resting-state functional connectivity (FC) between spatially segregated nodes of the DMN [27,28]. Increased connectivity and activation of the DMN could contribute to the ongoing self-referential processing and internal dialogue that maintains a hyperaroused state and perpetuates difficulties falling and remaining asleep [27,29]. Individuals with insomnia disorder show greater activation of the DMN compared with healthy controls while viewing word lists associated with past, present, and future worries, particularly when the words are self-referential [29].

Mainstream approaches to treating primary insomnia include cognitive behavioral therapy for insomnia and pharmacologic sleep aids, such as hypnotic sedatives [30-33]. While cognitive behavioral therapy for insomnia often helps to reduce ruminative cognitions and is effective at improving sleep for many individuals, there are many who fail to achieve meaningful benefits [34,35]. Similarly, pharmacologic sleep aids also tend to have modest effect sizes [36] and are often associated with unwanted side effects (eg, daytime drowsiness and memory problems) and health-related morbidities [37]. Various noninvasive neuromodulatory approaches have been shown to improve insomnia symptoms, including repetitive transcranial magnetic stimulation (rTMS) [38]. However, low-frequency rTMS has a cortical inhibition effect that has been shown to improve insomnia symptoms, which may be due to hyperarousal across multiple psychological and physiological domains [39].

The proposed protocol aims to bridge the gap between cognitive and neuromodulatory interventions to facilitate sleep onset and longer maintenance by temporarily inhibiting the brain FC that is associated with cortical hyperarousal and presleep ruminative cognitions. Our proposed approach will involve using rTMS prior to bedtime to briefly disrupt the strength of FC among cortical regions of the DMN. As we propose to inhibit connectivity within the DMN, we will use continuous theta burst stimulation (cTBS), which induces long-term depression of cortical neural firing following a sustained stimulation period of 40 seconds [40]. Thus, we hypothesize that the application of 40 seconds [40]. Thus, we hypothesize that the application...
of cTBS to an easily accessible node of the DMN will suppress the local neural activity of that node and propagate inhibition throughout the DMN, thereby reducing ruminative thinking and worry prior to sleep onset. Decreased DMN connectivity is expected to improve sleep quality and quantity. An increase in slow-wave and rapid eye movement sleep stages has been proposed as mechanisms of rTMS in increasing restorative sleep in individuals with insomnia, so polysomnography (PSG) parameters are the primary outcome measures in the outlined protocol [41]. This phase 1 clinical trial will be the first study to investigate the effects of cTBS targeted to the left inferior parietal node of the DMN on objectively measured sleep outcomes.

Research Aims and Hypotheses
This study aims to explore the effects of cTBS on (1) the activation and connectivity of the DMN and (2) sleep outcomes. We hypothesize that modulating the DMN by stimulating a targeted region of the left inferior parietal lobule with a single cTBS administration will decrease FC within the DMN and thereby improve sleep parameters (ie, PSG) relative to an identical sham administration.

Study Design and Randomization
This phase 1 clinical trial (ClinicalTrials.gov NCT04953559) will be conducted with a randomized, double-blind, sham-controlled, counterbalanced, crossover design. Participants serve as their own controls as they undergo 2 identical study sessions separated by at least 5 days (ie, washout period). Participants and study personnel who interact with the participants will be blind to the specific condition (active cTBS vs sham) the participant receives during each study visit. Using a prespecified, equally balanced, block randomization procedure, participants will be assigned to the order in which they receive active cTBS and sham stimulation conditions. The randomization list will be generated by a computer random number generator, with the constraint that half of the participants of each biological sex will receive sham first and half will receive the active cTBS first. This design will allow us to examine the intraindividual effects of the active cTBS and establish the superiority of active cTBS compared with sham stimulation in improving PSG sleep outcomes.

Methods

Study Procedures Overview
The study procedures were completed over 3 separate visits. Screening to determine eligibility and baseline assessments was conducted during visit 1. Eligible participants were asked to wear a wrist actigraph for at least 5 days prior to their first in-laboratory overnight visit (visit 2). Based on a preestablished counterbalanced crossover randomization schedule, participants were then assigned to receive either the active cTBS or sham stimulation for their first overnight visit (visit 2) followed by the alternative intervention during their second overnight visit (visit 3) at least 5 days later. As shown in Figure 1, each in-laboratory session lasted 18 hours, including a 2-hour cognitive testing block, 1 hour of preintervention magnetic resonance imaging (MRI), followed by either active cTBS or sham stimulation, and a postintervention hour of neuroimaging. Participants were then fitted with PSG electrodes and were permitted 8 hours of sleep while undergoing PSG sleep monitoring in the lab. The next morning, participants completed another 2-hour cognitive testing block before being released.

The full study execution takes approximately 38-40 hours with 2-3 hours allocated for baseline screening and enrollment (visit 1) and 18 hours allocated for each overnight visit (visits 2 and 3). Table 1 presents the specific assessment measures administered at each visit.

Figure 1. Overview of the pilot randomized controlled trial study design. Participants first complete a web-based eligibility survey screening for sleep disturbances. Potentially eligible volunteers complete an in-lab enrollment process and baseline assessment (visit 1) at a large Southwestern University medical research center. Participants are then randomized to a stimulation order condition. After at least 5 days with actigraphically measured sleep, participants return for an 18-hour overnight session involving 2 MRI scans and overnight sleep (visit 2) at the medical research center. Depending on their condition assignment order, they either receive active cTBS or sham intervention. Participants then undergo a washout period of at least 5 days that includes actigraphically measured sleep. They then return for an identical overnight visit that involves the alternate intervention condition. cTBS: continuous theta burst stimulation; MRI: magnetic resonance imaging.
Table 1. List of study activities and assessments performed during study visits during a pilot randomized controlled trial evaluating the preliminary effectiveness of continuous theta burst stimulation in improving sleep in individuals with reported sleep disturbances.

<table>
<thead>
<tr>
<th>Study activities and assessments</th>
<th>Baseline screening and enrollment (visit 1)</th>
<th>At home</th>
<th>Overnight visits (visits 2 and 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Screening</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consent form</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Eligibility screening (TMS(^a), health background, MRI(^b), PSQI(^c), ISI(^d), and ESS(^e))</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy test</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Covariates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Anxiety (STAI-S(^f) and STAI-T(^g))</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Depression (BDI-II(^h))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Sleep preoccupation (SPS(^i) and GCTI(^j))</td>
<td>√</td>
<td>✓</td>
<td>k</td>
</tr>
<tr>
<td>Caffeine consumption questionnaire</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Intelligence (WASI-II(^l))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Actigraphy (Phillips Actiwatch)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sleep quality (Sleep diaries)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Primary Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional connectivity in DMN(^m)</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Sleep (PSG(^n))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>Secondary Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information processing speed (GNG(^o))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Attention (PVT(^p))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Verbal memory (CVLT-3(^q))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Immediate and delayed memory (RBANS(^r))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Self-reported sleepiness (KSS(^s))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Mood (VAMS(^t))</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Side effects (pre- and post-TMS assessment)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

---

\(^a\)TMS: transcranial magnetic stimulation.
\(^b\)MRI: magnetic resonance imaging.
\(^c\)PSQI: Pittsburgh Sleep Quality Index.
\(^d\)ISI: Insomnia Severity Index.
\(^e\)ESS: Epworth Sleepiness Scale.
\(^f\)STAI-S: State-Trait Anxiety Inventory-State.
\(^g\)STAI-T: State-Trait Anxiety Inventory-Trait.
\(^h\)BDI-II: Beck Depression Inventory.
\(^i\)SPS: Sleep Preoccupation Scale.
\(^j\)GCTI: Glasgow Content of Thoughts Inventory.
\(^k\)Only GCTI.
\(^l\)WASI-II: Wechsler Abbreviated Scale of Intelligence.
\(^m\)DMN: default mode network.
\(^n\)PSG: polysomnography.
\(^o\)GNG: go/no-go task.
\(^p\)PVT: psychomotor vigilance task.
Ethical Considerations
All study activities and study personnel were approved by the University of Arizona’s Institutional Review Board (IRB) and by the Department of Defense Office of Human Research Oversight (OHRO) in March 2021 (approval 2007900971). Any protocol amendments and reportable new information were disseminated to both IRB and OHRO. All participant information was deidentified for confidentiality. Participants were compensated 500 dollars via check for their participation in the study if they were fully compliant with the study procedures. The study team maintains dissemination control of the final deidentified data set. Final results will only report cumulative population data to respective regulatory and scientific reporting agencies and along with peer-reviewed publications. Access to the deidentified data set will be available upon request to the principal investigator and in conjunction with proper regulatory requirement.

Participants and Recruitment
A total of 20 otherwise healthy adults with self-reported sleep disturbances were recruited from the greater Tucson and Phoenix metropolitan areas. Recruitment strategies include flyers posted on community boards, advertisements in local newspapers, sponsored social media posts, and email lists at local universities. Individuals who were recruited from the community were directed to complete a web-based survey in a HIPAA (Health Insurance Portability and Accountability Act)–compliant server. If participants met the initial inclusion criteria, a study member then completed a preliminary phone screening with the participant. Study personnel then scheduled an in-person screening visit for eligible individuals.

Power Analyses
During the study design, we conducted a power analysis to determine the sample size necessary to compare mean changes in DMN connectivity before and after cTBS using a within-subjects ANOVA approach. From prior cited work on transcranial magnetic stimulation (TMS) and FC, a total sample of 20 was found to have sufficient power to answer the research questions to include analysis of 2 groups (within-subject active and sham), 2 repeated measurements (before and after TMS), and a correlation between the change in DMN activation and sleep outcome metrics [42,43].

Inclusion and Exclusion Criteria
Individuals between the ages of 18-50 years with reported sleep disturbances were included in the study. Individuals were excluded if they reported or exhibited any health conditions beyond insomnia-related symptoms. Exclusionary criteria were based on one or more of the following considerations: the criterion in question is (1) known to alter sleep, (2) known to substantially increase interparticipant variability, (3) known to put the volunteer outside the range of what is considered healthy, or (4) required by regulation. Detailed inclusion and exclusion criteria are referenced in Textbox 1.
Textbox 1. A detailed description of inclusion and exclusion criteria of a pilot randomized controlled trial evaluating the preliminary effectiveness of continuous theta burst stimulation in improving sleep in individuals with reported sleep disturbances.

### Inclusion criteria
- 18-50 years old
- Self-reported sleep disturbances on 2 of 3 inventories:
  - Pittsburgh Sleep Quality Index ≥6
  - Insomnia Severity Index ≥15
  - Epworth Sleepiness Scale ≥11

### Exclusion criteria
- Unwillingness to provide informed consent
- Presence of a metal implant or medical device which poses a safety risk for magnetic resonance imaging or transcranial magnetic stimulation
- Self-reported past or present medical diagnosis of sleep- or breathing-related disorders
- Travel outside the time zone within 1 week prior to the enrollment visit or while active during the study
- Self-reported major medical or neurological problems
- Self-reported past or present history of cardiovascular disease
- Self-reported past or present history or first-degree family history of any seizures or seizure disorders
- Self-reported underlying acute or chronic pulmonary disease
- Self-reported history of fainting spells or syncope
- Self-reported past or present psychiatric problems
- Self-reported suicidal ideation
- Self-reported current use of prescription medications
- Self-reported current use of supplements that affect sleep
- Self-reported caffeine use in excess of 300 mg per day on average
- Self-reported use of cigarettes
- Self-reported use of illicit drugs
- Speaking English as a nonprimary language
- Less than a ninth-grade education
- Engaged in overnight shift work
- Female individuals only: positive urine pregnancy test
- Female individuals only: self-reported current breast-feeding or collecting breast milk

### Study Activities

#### Enrollment and Baseline Assessment (Visit 1)
Eligible participants were invited for an in-person screening and baseline data collection session lasting approximately 2-3 hours. After participant consent was obtained by trained study staff members, the participants underwent a baseline assessment, including measurement of intellectual capacity with the Wechsler Abbreviated Scale of Intelligence by a certified administrator, subjective sleep assessment including the Pittsburgh Sleep Quality Index and Insomnia Severity Index, and other cognitive assessments.

#### At-Home Actigraphy Week
Upon the completion of the enrollment visit, participants were fitted with a wrist actigraphic sleep monitor that they wear for at least 5 days prior to returning for the in-laboratory assessment sessions and continued to wear the device during the intervening washout week between the 2 in-residence laboratory stays. Participants also completed a web-based sleep diary each morning. Throughout the entire period of enrollment, participants were required to maintain a regular sleep schedule and were not permitted to use caffeine products for 48 hours prior to the in-laboratory overnight stays.
Overnight In-Residence Laboratory Sessions (Visits 2 and 3)

Participants completed 2 overnight in-residence laboratory stays according to a double-blind counterbalanced crossover design, with each session separated by at least 5 days. Figure 2 provides a graphical overview of the in-residence laboratory activities.

Figure 2. Timeline of the in-residence laboratory testing for the pilot randomized controlled trial. Participants arrive at the lab at 1500 and undergo cognitive testing followed by a preintervention MRI scan. The intervention is administered between 1900 and 2000 and is randomly assigned as either active cTBS stimulation or an identical appearing sham condition. After stimulation, participants complete a postintervention series of MRI scans, followed by a brief cognitive testing battery. At 2200, they are escorted to the sleep lab and PSG electrodes are applied. Lights out occur at 2300 and the participant is provided with an 8-hour undisturbed period for PSG-monitored sleep. Following wake-up at 0700 the next morning, the participant is provided a light breakfast and completes a final battery of cognitive tests. This procedure is repeated on separate weeks for the active cTBS and placebo conditions. cTBS: continuous theta burst stimulation; MRI: magnetic resonance imaging; PSG: polysomnography.

Cognitive Testing Block

Participants arrived at 3 PM and completed approximately 3 hours of cognitive testing with a trained and certified administrator, including the 10-minute psychomotor vigilance testing (PVT) and Karolinska Sleepiness Scale (KSS) at 3 time points separated by an hour, in addition to the Visual Analog Mood Scale, State-Trait Anxiety Inventory-State only (STAI-S), Beck Depression Inventory (BDI-II), California Verbal Learning Task 3 (CVLT3), go/no-go task (GNG), repeatable battery for the assessment of neuropsychological status (RBANS) symbol digit test, RBANS digit span, RBANS story memory test, and Glasgow Content of Thoughts Inventory (GCTI).

Prestimulation Neuroimaging

MRI scans were collected on a Siemens MAGNETOM Skyra 3T scanner (Siemens) using a 32-channel head coil. Participants completed a series of scans that included structural (magnetization-prepared rapid gradient-echo), functional (10-min eyes open) resting state, and proton magnetic resonance spectroscopy ($^1$H MRS) scans, with voxels placed in (1) the anterior cingulate gyrus and (2) posterior cingulate or precuneus. Functional MRI data will be preprocessed with standard neuroimaging packages, including statistical parametric mapping 12, FSL, and the functional connectivity toolbox (CONN v17f or later; focusing on DMN network connectivity) following standard published pipelines [44].

Intervention: cTBS Versus Sham

Following the first MRI scan, participants exited the scanner and underwent either cTBS stimulation or sham stimulation according to their preassigned double-blinded condition. The stimulator was preprogrammed prior to the start of the session and the participants and technicians that administered the stimulation procedure were blind to the administered condition. Each individual was expected to have a different sensitivity to the magnetic fields generated by the stimulation coil, and the stimulation intensity was adjusted based on each individual’s resting motor threshold (RMT). The real-time, motor-evoked potential of muscle contraction is provided to ensure consistent force production. Once the RMT was identified, the stimulation intensity for TBS was set to 70% of each participant’s RMT.

For primary rTMS stimulation, we located a predetermined node in the inferior parietal lobe, localized by coregistering the participant’s head, structural T1-weighted MRI image, and TMS coil in the same space using a TMS 3D Neuronavigation System. Once coregistration was complete, the Neuronavigation system provided real-time feedback on the TMS coil location and recorded the coil position and orientation relative to the head. The cTBS was applied using a figure-of-8 coil with an active cooling system connected to the MagPro magnetic stimulator (MagVenture Cool-B65). The same coil has 2 sides, 1 designed for active TMS stimulation, and the other is equipped with a magnetic shield that effectively blocks any stimulation and is used for the sham condition. The rTMS setup is shown in Figure 3.

For this project, we selected an easily accessible node of the DMN located on the left lateral parietal (LLP) cortex. The exact spatial location for stimulation is based on the Yeo et al [45] probability atlas. We downloaded this atlas into the Mango 4.1 visualization program (and identified the centroid of the LLP node of the DMN at the MNI coordinates of x ($-48$), y ($-61.5$), and z (32.5). These coordinates are converted into the same stereotaxic space as the participant’s brain to allow precision localization of the LLP node as the target for stimulation. Once this site was localized for the participants, a 40-second cTBS stimulation train (600 stimuli/session) or identically matched sham stimulation was administered using a MagVenture MagPro X100 stimulator (MagVenture Inc) connected with a figure-of-8 magnetic coil with active cooling.
Figure 3. Transcranial magnetic intervention system set-up for the pilot randomized controlled trial. Left: (A) the continuous theta burst stimulation is administered with a MagVenture Cool-B65 stimulator that included an active and sham side; (B) the resting motor threshold was determined using a MagPro X100 stimulator with a figure-of-8 coil; (C) the transcranial magnetic stimulation system is maintained at a constant temperature using an active cooling system. Middle: the stimulation is directed by a computerized system that detected the orientation of the head in space using (D) an antenna camera and correlate it with the individual’s coregistered magnetic resonance imaging scan. Right: An image showing the Neuronavigation system used to administer the transcranial magnetic stimulation to the precise default mode network locations.

Poststimulation Neuroimaging
Immediately following the cTBS or sham stimulation period, each participant underwent a second neuroimaging session that was identical to the prestimulation MRI session previously described.

PSG Recording
Participants slept undisturbed in a light- and temperature-controlled private bedroom while continuously monitored by a trained technician with a Nihon Kohden JE-921 PSG recording system using standard 10-20 electrode placement. Data from PSG were scored using Polysmith software according to the standard Rechtschaffen and Kales approach by a trained and certified sleep technician who was blind to the participant’s condition and status. The primary outcome metrics include sleep onset latency (SOL), latency to N1, N2, N3, and REM, total sleep time (TST), sleep efficiency (SE), wake after sleep onset (WASO), and number of awakenings and arousals, as well as time and percentage of time spent in each sleep stage.

Monitoring
Participants were assessed both prior to stimulation and after receiving stimulation for any adverse somatic, cognitive, or physical symptoms. Any presence of symptoms was logged in a HIPAA-compliant server and the participant was assessed by the study physician to determine if enrollment needed to be discontinued. All adverse events were reported to both IRB and OHRO. All study activity was subject to independent, random auditing from the Department of Psychiatry at the University of Arizona and study sponsors.

Data Management
All collected data were deidentified and stored within HIPAA-compliant, password-protected servers immediately after collection. Study data were only accessible to study personnel no more than 1 week after data collection. Data quality checks were conducted every quarter by staff members for additional data assurance. Any data collected after participants’ consent were included in the study database for both compliant and noncompliant participants. However, following a per-protocol analysis, we are only using data from participants who fully completed the study due to within-subjects analysis.

Primary Outcomes: FC Changes in the DMN
We hypothesize that active cTBS will disrupt the within-network and between-network connectivity of the DMN. This hypothesis will be assessed from a 10-minute resting-state FC scan and analyzed using standard procedures in the CONN (v17f or higher). Standard pipelines for preprocessing will be used. Within-network connectivity will be determined by first conducting an independent components analysis of the intrinsic activation patterns across the brain. Twenty independent components will be extracted. Mean activation clusters from each component will be examined for intercorrelation to determine within-network connectivity. Additionally, between-network, region-to-region, and seed-to-voxel analyses will be conducted to compare changes in FC between the 2 stimulation conditions.

Analysis Plan
FC Changes to the DMN
First-level whole brain connectivity analyses will be undertaken in CONN for each subject at each of the 4 sessions (ie, active prestimulation, active poststimulation, sham prestimulation, and sham poststimulation). This will result in an FC map for each participant at each session, indicating the magnitude and direction of intrinsic correlation between each seed region (eg, lateral parietal cortex of the DMN) and all other voxels in the brain. For our primary analysis, we will estimate the effects of active cTBS on FC within the DMN using multimodal neuroimaging data including resting state FC. Raw NifTI images
will be preprocessed using standard functional preprocessing pipelines in CONN [46].

The first-level connectivity maps will be combined in subsequent second-level analyses in CONN to explore the effects of cTBS stimulation on brain connectivity to include (1) independent components analysis, (2) seed-to-voxel analyses, and (3) region of interest to region of interest analyses [44]. Statistically significant differences in FC of the DMN between sham and active conditions will be the minimal important change to confirm the primary hypothesis. Next, PSG parameters measuring sleep stage duration and onsets, arousals, and awakenings will be added as covariates of FC change from before to after stimulation, comparing sham and active cTBS conditions.

**PSG Analysis**

We will assess the effects of active cTBS on sleep and whether participants demonstrated greater sleep quality during their active cTBS session compared with their sham stimulation session. Analyses will be conducted on sleep variables including total sleep time; N1 latency; N2 latency; N3 latency; REM latency; SOL; persistent SOL; SE; WASO; total number of awakenings; spontaneous arousals; wake duration; and stage duration for N1, N2, N3, and REM. Analyses will include repeated measures ANOVA controlling for covariates. Statistically significant difference between sham and active conditions will be the minimal important change to confirm this hypothesis.

**Exploratory Analysis**

We will examine the effects of active cTBS on secondary outcomes of cognitive performance, mood, and side effects compared with sham stimulation using repeated measures ANOVAs and nonparametric tests accordingly.

**Results**

Our study was funded in September 2020 by the Department of Defense (W81XWH2010173). We completed the enrollment of our target population of 20 participants in October 2022. As of July 2023, we will initiate extensive neuroimaging data analysis and anticipate that the full results of the study will be published by 2024.

**Discussion**

**Study Rationale**

With a growing number of people reporting greater sleep disturbances, identifying effective nonpharmacological interventions is an important step in providing more innovative and noninvasive approaches for improving sleep outcomes. While the sample size for this study is small, this is the first study to investigate cTBS effects on the DMN to improve sleep outcomes.

**Experimental Clinical Design Rationale**

Double-blind randomized controlled trials are the gold standard for scientific inquiry. Therefore, each participant will serve as their own control in a counterbalanced crossover design. Implementing a crossover design in which every participant receives both active and sham cTBS stimulation permits comparison of intraindividual differences in DMN connectivity and sleep quality parameters, as each participant acts as their own control. Because the goal was to examine the effects of cTBS on individuals with sleep problems, no healthy control samples were recruited or necessary to test our hypotheses. Blinding study personnel to the TMS condition also reduces systematic differences in the administration of the stimulation interventions and ostensibly reduces the influence of expectations about treatment effects.

**Brief cTBS of the LLP Lobule**

Continuous theta burst was selected as the intervention based on its previously demonstrated inhibitory effects on the cortex [47]. A single 40-second cTBS session inhibits, or suppresses, neural excitability in the cortex for up to 50 minutes or longer [47]. Stimulation of a single node of the DMN is expected to inhibit local activity, with further suppression of cortical excitability propagating throughout the network overall. Inhibition of the left inferior parietal lobule was selected for two primary reasons: (1) prior research showed reduced metabolic activity in this region when individuals were falling asleep and (2) this region facilitates an easy-to-access node of the posterior DMN [48]. The localization of specific brain regions requires prestimulation imaging to account for individual variability in brain structure. For this reason, a critical element of that project was to show that stimulation of a single, easily accessible surface node of the larger DMN could lead to significant alterations in FC within this network. Prior investigations using cTBS have focused primarily on the stimulation of brain regions associated with the prefrontal areas of the DMN. Sleep-related parameters have rarely been collected as outcome measures for cTBS. One prior study examining rTMS effects on the right parietal cortex showed reduced anxiety and insomnia symptoms [49]. However, that study used a stimulation session (of three 1-Hz stimulations every 10 minutes) spread across 10 days. The presently proposed cTBS protocol is aimed at establishing initial research outcomes of the effects of a single burst of cTBS on disrupting network connectivity of the DMN to improve sleep. Thus, only a single stimulation session of both active cTBS and sham is needed to identify the intraindividual effects of cTBS on sleep outcomes.

**Limitations and Future Directions**

Targeting a single node of the DMN and conducting only 1 stimulation session are limitations of this protocol to sufficiently answer the research questions and aims proposed regarding DMN and insomnia improvement. Future work could determine the most effective stimulation sites within the DMN and the optimal number of pulses and administrations to optimize this potential intervention for sleep problems. Larger samples are also needed than the small pilot sample collected for this protocol to examine the effects of sex observed in previous studies [38].

**Summary and Conclusions**

Insomnia symptoms linked with many other psychiatric and physiological morbidities have become an increasingly prevalent
and important health care and societal concern. Noninvasive neuromodulation can be inexpensive and is a technique accessible to many types of treatment providers. The protocol presented here unifies the prominent aspects of insomnia—physiological and psychological hyperarousal—by targeting a unique neural network, known to be associated with proposed mechanisms of insomnia [8,50].

Acknowledgments

This study was funded by the United States Army Medical Research and Development Command (W81XWH2010173; primary investigator: WDSK).

Authors' Contributions

LH carried out the data collection, data management, wrote the initial draft of the manuscript, and contributed to revisions. AH contributed equally to the initial drafting of the manuscript, provided revisions, and contributed to the ongoing statistical analysis of the study. ND assisted with the study design and executed data collection. SJ and KHA contributed to data management, and regulatory reporting, and reviewed drafts of the manuscript. CT, YCC, and YHC provided training in transcranial magnetic stimulation (TMS), provided the TMS study equipment, and oversaw TMS data collection. WDSK was responsible for the initial conceptualization and design of the project, obtaining the research funding, providing oversight of the study, contributing to ongoing statistical analysis and interpretation of the findings, and contributing to initial drafting and revisions of the manuscript.

Conflicts of Interest

None declared.

References


Abbreviations

- **ANOVA**: analysis of covariance
- **BDI-II**: Beck Depression Inventory
- **CONN**: Functional Connectivity Toolbox
- **cTBS**: continuous theta burst stimulation
- **CVLT-3**: California Verbal Learning Task 3
- **DMN**: default mode network
- **FC**: functional connectivity
- **GCTI**: Glasgow Content of Thoughts Inventory
- **HIPAA**: Health Insurance Portability and Accountability Act
- **IRB**: Institutional Review Board
- **LLP**: left lateral parietal
- **MRI**: magnetic resonance imaging
- **OHRO**: Office of Human Research Oversight
- **PSG**: polysomnography
- **PVT**: psychomotor vigilance task
- **RBANS**: Repeatable Battery for the Assessment of Neuropsychological Status
- **RMT**: resting motor threshold
rTMS: repetitive transcranial magnetic stimulation
SE: sleep efficiency
SOL: sleep onset latency
TMS: transcranial magnetic stimulation
TST: total sleep time
WASO: wake after sleep onset
Epidemiology of Syphilis in Pregnancy and Congenital Syphilis in Brazil and the Risk or Associated Factors: Protocol for a Systematic Review

Yago Tavares Pinheiro¹, PhD; Janmilli da Costa Dantas¹, PhD; Jose Rebbery Rodrigo Holanda², MD; Ankilma do Nascimento Andrade Feitosa³, PhD; Richardson Augusto Rosendo da Silva¹, PhD

¹Federal University of Rio Grande do Norte, Natal, Brazil
²Federal University of Rio Grande do Norte, Caicó, Brazil
³University Center Santa Maria, Cajazeiras, Brazil

Corresponding Author:
Yago Tavares Pinheiro, PhD
Federal University of Rio Grande do Norte
Lagoa Nova
Natal, 59078-900
Brazil
Phone: 55 (83) 99664 7797
Fax: 55 (83) 99664 7797
Email: yagosta5@gmail.com

Abstract

Background: Syphilis in pregnancy and congenital syphilis are growing public health issues worldwide. Several factors can influence their occurrence in the population. Therefore, understanding the epidemiology of this condition and the factors that influence its occurrence is fundamental for decision-making by clinicians and health managers. However, so far, no systematic review has summarized and analyzed data on the incidence, prevalence, and predictors of these diseases in Brazilian cities, considering different sociocultural, demographic, economic, sanitary, and spatial-temporal characteristics presented across locations.

Objective: We propose a systematic review protocol to gather and analyze data on the incidence, prevalence, and risk or associated factors of syphilis in pregnancy and congenital syphilis in Brazil, taking into account different local or regional contexts.

Methods: Searches will be conducted in CINAHL, MEDLINE, LILACS, Embase, and Web of Science databases. We will include observational studies (ie, cross-sectional, longitudinal, or case-control studies), analyzing the incidence, prevalence, and predictors of these diseases in Brazilian cities, considering different sociocultural, demographic, economic, sanitary, and spatial-temporal characteristics presented across locations.

Results will be discussed based on subgroup analysis, which is as follows: (1) type of syphilis (in pregnancy or congenital), (2) type of study (case-control and cross-sectional studies for analysis of associated factors and longitudinal studies for risk factors), and (3) contextual factors (ie, region of country, socioeconomic and demographic characteristics, and year of study). This systematic review is expected to be completed by December 2023, and our results will be disseminated through publication in peer-reviewed journals and scientific events.

Conclusions: This systematic review aims to assist health care managers and professionals in their decision-making to control these diseases in Brazil, considering location heterogeneity. Furthermore, countries with health systems and demographic and socioeconomic contexts similar to those of Brazil may benefit from this information.
Introduction

Syphilis is a sexually transmitted infection caused by *Treponema pallidum* and is a major public health issue worldwide [1]. Despite the efforts of health care professionals to control this infection in Brazil, cases of syphilis have increased in recent years [2-4], impacting public and private health care systems and highlighting the need to improve disease surveillance [5,6]. Globally, 2 million out of 36 million syphilis infections occur in pregnant women [7], resulting in congenital syphilis (infection of the fetus) and adverse events (eg, early fetal death, stillbirth, premature birth, low birth weight, and neonatal death) [8,9].

Recently, an outbreak of syphilis has been observed among men and women in more economically developed countries, which can be explained by changes in the sexual behavior of individuals and increased exposure to the risk of infection due to a false sense of security stemming from new treatments and an increased search for sexual partners over the internet [10]. In this context, understanding the epidemiology and control of this disease becomes more complex and difficult.

Syphilis in pregnancy and congenital syphilis can be controlled with health care measures, such as access to prevention services, early diagnosis, and treatment [4,11,12]. Conversely, these measures require the analysis of epidemiological data and predictors [4,11]. Although studies in Brazilian cities analyzed the incidence, prevalence, and predictors of syphilis in pregnancy and congenital syphilis [13-17], each city presented different sociocultural, demographic, economic, sanitary, and spatial-temporal characteristics, hindering data extrapolation to the national territory.

Summarizing and analyzing the incidence, prevalence, and risk or associated factors of syphilis in pregnancy and congenital syphilis must take into account location heterogeneity. This type of analysis enables a broader understanding of the problem, improves control strategies and equity in disease management, and establishes reference data to help disease screening efforts in Brazil. Despite the relevance of the theme, no systematic review has been conducted to date on the epidemiological data of gestational and congenital syphilis or its predictors, subgrouping and analyzing this information from different contexts.

Thus, we propose a systematic review protocol to gather and analyze data on the incidence, prevalence, and risk or associated factors of syphilis in pregnancy and congenital syphilis in Brazil, taking into account different local or regional contexts.

Methods

Study Design

This systematic review protocol was developed according to PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) [18], which will also guide the systematic review. The protocol has been registered in PROSPERO (CRD42022329329).

Eligibility Criteria

Observational studies whose sample comprised cases of syphilis in pregnant women or newborns in Brazil will be included in the systematic review. Table 1 [2,19,20] presents the eligibility criteria used in the review.
Table 1. Eligibility criteria for the systematic review.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design</td>
<td>• Cross-sectional, longitudinal, or case-control studies conducted in Brazil</td>
<td>• Reviews, opinion articles, editorials, or publications without primary data or not peer-reviewed</td>
</tr>
<tr>
<td></td>
<td>• Studies based on primary data</td>
<td>• The most complete and recent data will be used if studies report the same data in multiple sources [19,20].</td>
</tr>
<tr>
<td>Population and</td>
<td>• Studies in which the sample involved pregnant women or new borns with syphilis</td>
<td>• Studies with samples involving other populations (non-pregnant women or men)</td>
</tr>
<tr>
<td>location</td>
<td>• Studies with residents in Brazil</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>• Studies reporting the prevalence or incidence of Treponema pallidum (syphilis)</td>
<td>• Studies presenting the incidence or prevalence of combined infections (ie, syphilis with other sexually transmitted infections) and not allowing isolated analysis</td>
</tr>
<tr>
<td></td>
<td>infection or its risk or associated factors in pregnant women or newborns</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Studies that diagnosed syphilis based on direct pathogen detection tests or</td>
<td></td>
</tr>
<tr>
<td></td>
<td>through immunological, treponemal, or non-treponemal tests, in accordance with</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brazilian protocols for diagnosing syphilis [2]</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>—a</td>
<td>• No restriction regarding language or year of publication</td>
</tr>
<tr>
<td>Time frame</td>
<td>—</td>
<td>• No restriction regarding to year in which the study was carried out or published</td>
</tr>
</tbody>
</table>

aNot applicable.

Study Selection

Textbox 1 presents the search strategy. Searches will be conducted in the CINAHL, MEDLINE, LILACS, Embase, and Web of Science databases. Grey literature will be also searched using the reference lists of relevant studies in addition to using databases such as Open Gray and Google Scholar. Further, reports with epidemiological data on syphilis in the country will be screened in the electronic database of the Brazilian Ministry of Health.

Two researchers will independently search, identify potentially eligible studies, and remove duplicates. Then, inclusion and exclusion criteria will be applied to titles and abstracts, eligible studies will be read in full, and reasons for exclusion will be recorded. Disagreements between researchers will be resolved by discussion or with a third researcher. The flowchart of the study selection is described in Figure 1.

Textbox 1. Search strategy.

```
#1 "syphilis" [Title/Abstract] OR "congenital syphilis" [Title/Abstract] OR “treponemal infections” [Title/Abstract] OR “T. pallidum” [Title/Abstract] OR "pallidum" [Title/Abstract] OR "serosyphilis" [Title/Abstract] OR “sexually transmitted diseases” [Title/Abstract]
#2 "pregnant" [Title/Abstract] OR “women” [Title/Abstract] OR “congenital” [Title/Abstract]
#3 “incidence” [Title/Abstract] OR “prevalence” [Title/Abstract] OR “prevalence study” [Title/Abstract] OR “cross-sectional study” [Title/Abstract] OR “observational study” [Title/Abstract]
#4 “risk factors” [Title/Abstract] OR “associated factors” [Title/Abstract] OR “measures of association, exposure, risk or outcome” [Title/Abstract]
#5 “brazilian” [Title/Abstract] OR “brazilian” [Title/Abstract]
#6 #1 AND #2 AND #3 OR #4 AND #5
```
Assessment of Methodological Quality and Quality of Evidence

The Newcastle-Ottawa Scale will be used to assess the methodological quality of studies [21]. This scale includes 8 items categorized into 3 domains (ie, selection, comparability, and outcome or exposure) to assess the risk of bias in nonrandomized studies. The Newcastle-Ottawa Scale has specific tools for cohort and case-control studies. Thus, adaptations will be made to allow the proper assessment of the potential sources of bias in cross-sectional studies. In addition, quality of evidence will be analyzed using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluations) system, which classifies evidence as high, moderate, low, and very low [22,23].

Data Extraction and Synthesis

The following data will be extracted and entered into a Microsoft Excel spreadsheet: first author, year of publication, type of study (eg, longitudinal, cross-sectional, or case-control), type of syphilis (eg, in pregnancy or congenital), the diagnostic method used, study location (eg, city and state), participants (eg, sample size, age, type of population, and presence of co-infection), population setting (eg, community, health care centers, schools, neighborhoods, and the environmental context of the participants), date of data collection, sampling method, and main results (eg, incidence, prevalence, and risk or associated factors).

After summarizing the studies, results will be discussed based on subgroup analysis, as follows: (1) type of syphilis (eg, in pregnancy or congenital), (2) type of study (eg, case-control and cross-sectional for analysis of associated factors and longitudinal for risk factors), and (3) contextual factors (eg, region of country, socioeconomic and demographic characteristics, and year of study).

Statistical Planning

Kappa coefficient will assess the agreement between researchers [24]. The unadjusted incidence or prevalence and the standard error will be recalculated based on the numerator and denominator values presented in each study. Furthermore, the prevalence or incidence may be reported using the direct method of standardization, adjusted for the variables of age, study location, and presence of co-infection. If the study does not provide data for calculating adjusted incidence or prevalence, the researchers will request this information from the study authors.

Additionally, a meta-analysis will be performed using a random-effects model due to the potential heterogeneity among studies. The random-effects model is applied when the aim is to combine several studies that have similar objectives but are conducted in different ways (ie, exhibiting methodological
heterogeneity) [25]. Moreover, the Freeman-Tukey double arcsine transformation will stabilize variances to maintain the estimates of individual effects of each study [26]. Cochran Q test will assess the heterogeneity among studies [27]. $I^2$ values of 25%, 50%, and 75% will represent low, medium, and high heterogeneity, respectively [28].

Studies will undergo a subgroup analysis using clustering variables (eg, study location, study population, method of syphilis diagnosis, mean sample size, year of data collection, sampling methods, and methodological quality) to investigate possible sources of heterogeneity [20].

Analyzes will be performed using the Review Manager (RevMan) software (version 5.4; Cochrane Collaboration) and the R software (R Core Team), considering a 95% CI.

Results

The protocol has been registered in PROSPERO (CRD42022329329). The screening of the studies in the databases has already started, and the entire systematic review is expected to be completed by December 2023. The results of the study will provide evidence that can support decision-making regarding strategies to control syphilis in Brazil and countries with similar health, demographic, and socioeconomic profiles.

Results will be disseminated through publication in peer-reviewed journals and presentation at scientific events.

Discussion

Expected Results and Practical Implications

After a preliminary search, we found studies in Brazilian cities that analyzed epidemiological data and predictors of syphilis in pregnancy and congenital syphilis [13-17,29]. However, no study has organized and summarized data to perform a broader analysis of this public health issue.

Summarizing local studies will allow the analysis and discussion of epidemiology and risk or associated factors of syphilis in pregnancy and congenital syphilis, considering sociocultural, demographic, spatial-temporal, economic, and sanitary differences in each location. Thus, this systematic review will help in the decision-making of health care managers and professionals to control these diseases in Brazil according to location heterogeneity.

Limitations

Some limitations that may compromise the quality of evidence can be found in the systematic review, such as heterogeneity among studies, wide CIs, and uncertainty of estimated effects.

Acknowledgments

The authors would like to thank Probatus Academic Services for providing scientific language translation and revision.

This study was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil (finance code 001).

Data Availability

All data generated and analyzed in this study are available upon request from the corresponding author.

Authors’ Contributions

YTP, JCD, and ANAF designed the study and wrote the original draft; JRRH and RARS designed the study and approved the final manuscript.

Conflicts of Interest

None declared.

References


Abbreviations

GRADE: Grading of Recommendations, Assessment, Development, and Evaluations
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols

Edited by A Mavragani; submitted 10.07.23; peer-reviewed by C Shah; comments to author 11.10.23; revised version received 13.10.23; accepted 14.10.23; published 04.01.24.

Please cite as:
Pinheiro YT, Dantas JDC, Holanda JRR, Feitosa ADNA, Augusto Rosendo da Silva R
Epidemiology of Syphilis in Pregnancy and Congenital Syphilis in Brazil and the Risk or Associated Factors: Protocol for a Systematic Review
JMIR Res Protoc 2024;13:e50702
URL: https://www.researchprotocols.org/2024/1/e50702
doi:10.2196/50702
PMID:38175689

©Yago Tavares Pinheiro, Janmilli da Costa Dantas, Jose Rebbery Rodrigo Holanda, Ankilma do Nascimento Andrade Feitosa, Richardson Augusto Rosendo da Silva. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 04.01.2024. 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

Reward Learning as a Potential Mechanism for Improvement in Schizophrenia Spectrum Disorders Following Cognitive Remediation: Protocol for a Clinical, Nonrandomized, Pre-Post Pilot Study

Frances Dark¹,², MBBS, PhD; Graham Galloway³,⁴, PhD, BSc(H); Marcus Gray³, PGDipPsy, PhD; Matteo Cella⁵, MScClinPsych, PhD; Veronica De Monte¹, MClinPsych, PhD; Victoria Gore-Jones¹, BS, BSocSci, MClinPsych; Gabrielle Ritchie¹,², MClinPsych, PhD

¹Metro South Addiction and Mental Health Services, Woolloongabba, Australia
²Faculty of Medicine, University of Queensland, Brisbane, Australia
³Translational Research Institute, Woolloongabba, Australia
⁴Herston Imaging Research Facility, The University of Queensland, Brisbane, Australia
⁵King's College London, London, United Kingdom

Corresponding Author:
Frances Dark, MBBS, PhD
Metro South Addiction and Mental Health Services
Level 2 Woolloongabba Community Health Centre
228 Logan Road
Woolloongabba, 4102
Australia
Phone: 61 0733171080
Email: frances.dark@health.qld.gov.au

Abstract

Background: Cognitive impairment is common with schizophrenia spectrum disorders. Cognitive remediation (CR) is effective in improving global cognition, but not all individuals benefit from this type of intervention. A better understanding of the potential mechanism of action of CR is needed. One proposed mechanism is reward learning (RL), the cognitive processes responsible for adapting behavior following positive or negative feedback. It is proposed that the structure of CR enhances RL and motivation to engage in increasingly challenging tasks, and this is a potential mechanism by which CR improves cognitive functioning in schizophrenia.

Objective: Our primary objective is to examine reward processing in individuals with schizophrenia before and after completing CR and to compare this with a group of matched clinical controls. We will assess whether RL mediates the relationship between CR and improved cognitive function and reduced negative symptoms. Potential differences in social RL and nonsocial RL in individuals with schizophrenia will also be investigated and compared with a healthy matched control group.

Methods: We propose a clinical, nonrandomized, pre-post pilot study comparing the impact of CR on RL and neurocognitive outcomes. The study will use a combination of objective and subjective measures to assess neurocognitive, psychiatric symptoms, and neurophysiological domains. A total of 40 individuals with schizophrenia spectrum disorders (aged 18-35 years) will receive 12 weeks of CR therapy (n=20) or treatment as usual (n=20). Reward processing will be evaluated using a reinforcement learning task with 2 conditions (social reward vs nonsocial reward) at baseline and the 12-week follow-up. Functional magnetic resonance imaging responses will be measured during this task. To validate the reinforcement learning task, RL will also be assessed in 20 healthy controls, matched for age, sex, and premorbid functioning. Mixed-factorial ANOVAs will be conducted to evaluate treatment group differences. For the functional magnetic resonance imaging analysis, computational modeling will allow the estimation of learning parameters at each point in time, during each task condition, for each participant. We will use a variational Bayesian framework to measure how learning occurred during the experimental task and the subprocesses that underlie this learning. Second-level group analyses will examine how learning in patients differs from that observed in control participants and how CR alters learning efficiency and the underlying neural activity.
**Results:** As of September 2023, this study has enrolled 15 participants in the CR group, 1 participant in the treatment-as-usual group, and 11 participants in the healthy control group. Recruitment is expected to be completed by September 2024. Data analysis is expected to be completed and published in early 2025.

**Conclusions:** The results of this study will contribute to the knowledge of CR and RL processes in severe mental illness and the understanding of the systems that impact negative symptoms and cognitive impairments within this population.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/52505

---

**KEYWORDS**
cognitive remediation; fMRI; functional magnetic resonance imaging; negative symptoms; psychosocial functioning; reward learning

---

**Introduction**

**Overview**

Schizophrenia spectrum disorders (SSDs; eg, schizophrenia, schizoaffective disorder, and schizophreniform disorder) are defined by the presentation of a range of symptoms, including delusions, hallucinations, disorganized thoughts and behaviors, and negative symptoms (ie, reduced motivation and emotional expression) [1]. These disorders are linked to a variety of deficits in cognition, which extend to both neurocognition (eg, attention, memory, and planning) and social cognition (eg, difficulties perceiving and processing emotions) [2]. Both stand-alone and integrated programs have been used to treat these cognitive impairments. Cognitive remediation (CR) therapy has been widely used as an intervention for deficits in global cognition and functional difficulties in schizophrenia and has been shown to be particularly beneficial if rehabilitation is also incorporated into treatment [3,4]. Nevertheless, response to therapy is variable [5], and the impact individual characteristics have on the success of CR is still being investigated. The mechanisms by which CR is effective are yet to be clarified. A core component of CR is strategic learning principles, which ensure tasks are scaffolded based on previous successful achievements and the chances of successful task completion are optimized. There is, therefore, close reinforcement of learning.

One proposed mechanism for the effect of CR is reward learning (RL), which is potentially the pathway to improved cognition and motivational negative symptoms. In schizophrenia, there is impairment in reward anticipation [6,7] and representation [8], leading to poorer decision-making and goal-directed behavior, motivational deficits, and negative symptoms [8-11]. RL is a term used to identify the cognitive processes responsible for adapting behavior following positive or negative feedback. RL is a basic adaptive function of every living organism and provides the possibility to adapt and change in response to internal and environmental demands [12]. This process has been extensively studied in neuroscience and linked to the brain dopamine system. The dopamine hypothesis of schizophrenia is the single most influential theory in our understanding of the neurochemical basis of the disorder [13]. This theory suggests that fundamental dysregulation in this system is responsible for the illness’s symptoms. Dysregulation in the dopamine system is also linked to RL abnormalities, which, in turn, are thought to influence cognitive and negative symptoms. A growing body of basic neuroscience literature has identified 2 complementary and interactive neural systems in the dopamine system responsible for predicting outcomes and learning from feedback [14]. The first of these systems, responsible for rapid learning, is mediated by the basal ganglia. This system, referred to as the “fast system,” is believed to represent the predicted value of actions and rewards. These predictions bias actions and underlie learning based on positive and negative feedback. The second slower system is based primarily in the prefrontal cortex and allows for more detailed, conscious, and abstract representations of values and rewards. These representations of value are instrumental in allowing individuals to flexibly respond to reward value and adapt to novelty in the environment.

There is consistent evidence that people with SSDs are impaired at making rapid behavioral adjustments in response to feedback and that these impairments are associated with negative and cognitive symptoms [15-17]. Problems using the “fast system” are evident in situations requiring rapid change in responses to environmental changes when a situation previously rewarding begins to be associated with disadvantageous outcomes and oversensitivity to negative feedback and poor sensitivity to positive feedback. In contrast, several studies suggest that the gradual or procedural learning system seems intact in people with schizophrenia [18], but antipsychotic medication dosage, particularly for those with high levels of dopamine 2 receptor blockades, may exert a negative effect on this system.

Social environments are dynamic with constant rapid changes; hence, social situations require rapid behavioral adjustments in response to ever-changing social feedback. People with SSDs have impaired social functioning, and recent studies have shown that they also have impaired social reward processing [19,20]. Social approval induces rewarding feelings and is associated with increased activation in regions and networks associated with reward [21-23]. In those with SSDs, there is reduced activity in common reward brain regions during the experience of social reward [24], suggesting that they may have a reduced experience of the rewarding feeling of positive social attention. Positive social interactions have benefits for mental well-being and give life a sense of meaning [25]. Receiving praise and attention from others improves self-esteem [26] and increases motivation [27]. Although social reward has major impacts on functional outcomes, only recent efforts have explored social reward processing in SSDs. Further behavioral evidence suggests that RL difficulties are more pronounced in learning from positive, rather than negative, feedback [8]. This provides a further link between the effects of impaired RL and negative
symptoms. Learning preferentially from negative outcomes is likely to lead to behavioral avoidance and social withdrawal, and have a negative impact on motivation. This hypothesis is supported by research suggesting that the magnitude of RL impairment, particularly for positive feedback, is associated with negative symptom severity [28].

Despite the significance of RL problems in people with SSDs, there is no therapy targeting this problem. One previous study explored the impact that a course of CR has on RL problems in people with schizophrenia [29]. The results of this study showed that CR could improve sensitivity to positive and negative feedback and that improvement in these parameters was moderated by the severity of negative symptoms. However, this study used a standard CR protocol and may not have achieved the maximum effect on RL problems. Furthermore, the nature of this study did not allow for investigating the retention of RL improvements and, more crucially, how these may impact cognitive and negative symptoms and, more broadly, recovery. RL difficulties in people with SSDs are associated with negative symptoms, and it is plausible that, by reducing RL difficulties, a reduction in negative symptoms could be observed. It is proposed that the structure of CR enhances rewards and motivation to engage in increasingly challenging tasks, and this is a potential mechanism by which CR can achieve functional outcomes in individuals with SSDs.

**Aims and Hypotheses**

Our primary aim is to investigate reward processing in individuals with SSDs before and after completing a course of CR and to compare this intervention group with a treatment-as-usual (TAU) group of individuals with SSDs. In addition, this study aims to investigate whether RL mediates changes in cognitive function and negative symptoms following CR. We will also examine potential differences in social RL compared with nonsocial RL in individuals with SSDs and the impact of CR on these potential differences in RL domains. Comparison with RL in a healthy adult control group will allow further differentiation of behavioral and neural impairments in SSDs.

**Hypothesis 1A**

At baseline, participants with SSDs will demonstrate deficits in RL when compared with healthy control volunteers. These differences in learning will be linked to aberrant activity in the dopamine system at a neural level in the prefrontal cortex and subcortical structures such as the basal ganglia, compared with healthy controls.

**Hypothesis 1B**

At baseline, participants with SSDs will demonstrate greater deficits in social RL when compared with nonsocial RL.

**Hypothesis 2**

Participants that complete CR will demonstrate improved RL, again reflecting improved neural activity within the prefrontal and basal ganglia regions, when compared with people with SSDs not receiving the intervention.

**Hypothesis 3**

Reward processing ability will mediate improved cognition following CR in participants receiving the intervention but not in the SSD control group.

**Hypothesis 4**

Reward processing will mediate improvements in negative symptoms following CR in participants receiving the intervention but not in the SSD control group.

**Methods**

**Study Design**

The study is a nonrandomized clinical pilot trial to investigate whether reward processing pathways are involved in the mechanism of action of cognitive remediation therapy (Computerized Interactive Remediation of Cognition-Training for Schizophrenia [CIRCuiTS]) [30,31]. We will recruit 3 participant groups: 20 participants with SSDs who will complete the CR program (the intervention group), 20 participants with SSDs who will receive TAU, and 20 matched healthy control participants. Figure 1 depicts a schematic of the study design.
Participants

A total of 40 participants with a diagnosis of an SSD will be recruited from the Metro South Addiction and Mental Health Service community teams (Brisbane, Australia). Given that this is a pilot study, a power analysis was not conducted [32]. Instead, the sample size was informed by practical factors relating to the project, including the budget, the availability of CR facilitators, and the recruitment and attrition rates from our previous studies. This sample size is also consistent with other research in this area [33].
The inclusion criteria for the intervention and TAU groups and the healthy control group, as well as the exclusion criteria for all groups, have been provided in Textbox 1.

Textbox 1. The inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria for the intervention and treatment-as-usual (TAU) groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Patients of the Metro South Addiction and Mental Health Service</td>
</tr>
<tr>
<td>• Aged between 18 and 35 years</td>
</tr>
<tr>
<td>• Primary diagnosis of schizophrenia spectrum disorder (SSD)</td>
</tr>
<tr>
<td>• English literacy skills of at least grade 4 equivalence</td>
</tr>
<tr>
<td>• Absence of neurological disorders or acquired brain injury</td>
</tr>
<tr>
<td>• Estimated intelligence quotient &gt;70 on the Test of Premorbid Functioning (TOPF)</td>
</tr>
<tr>
<td>• Agree to participate and have the capacity to consent to the study procedures</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inclusion criteria for the healthy control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Aged between 18 and 35 years</td>
</tr>
<tr>
<td>• No history of diagnosis of SSD</td>
</tr>
<tr>
<td>• English literacy skills of at least grade 4 equivalence</td>
</tr>
<tr>
<td>• Absence of neurological disorders or acquired brain injury</td>
</tr>
<tr>
<td>• Estimated intelligence quotient &gt;70 on the TOPF</td>
</tr>
<tr>
<td>• Agree to participate and have the capacity to consent to the study procedures</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria for all groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Metallic object in their body (eg, cardiac pacemaker and cochlear implant)</td>
</tr>
<tr>
<td>• Pregnant or possibly pregnant (unprotected sex since last menstrual period)</td>
</tr>
<tr>
<td>• History of claustrophobia</td>
</tr>
<tr>
<td>• Permanent metal dental appliances</td>
</tr>
<tr>
<td>• Bodyweight ≥120kg</td>
</tr>
</tbody>
</table>

The researchers will obtain consent from all participants through a participation information and consent form. For participants with SSDs, half (n=20) will complete CR and form the intervention group. Those participants with SSDs (n=20) that are not interested in completing CR will form the patient TAU group and will complete the pre- and postmeasures only. Thus, group allocation will be based on an individual’s preference to participate in CR. In addition, matched healthy controls (n=20) will be recruited from the Metro South Addiction and Mental Health community services in the Princess Alexandra Hospital district as well as from the general population. We do not anticipate that it will be difficult to collect these healthy controls, given our already established connections with other researchers and staff within this district. Therefore, we expect to be able to recruit this group through word of mouth and the snowball effect. These health controls will have no history of SSDs. This will provide a benchmark to compare the clinical groups. Hence, the total number of participants for the study is 60.

Intervention

CIRCuiTS is a therapist-supported CR web-based program that focuses on improving cognitive skills, particularly for individuals with psychosis [30,31]. Participants work through different cognitive tasks and exercises, many of which are based on real-life experiences (eg, creating a shopping list or cooking).

Previous studies have shown CIRCuiTS leads to improvements in both cognition and functional recovery and is acceptable by participants [34]. Sessions typically last 1 hour, twice a week, for 12 weeks, and task practice is delivered through a computer. CIRCuiTS consists of 40 stages. A total of 20 sessions are considered adequate treatment exposure, and 20 minutes is the minimum time for a “session.” In this study, participants in the CR group will have 1 face-to-face meeting with the therapist to orient themselves to the program. These participants will then complete the program either on the web or in person, in a group or individually. Only the intervention group will be able to complete these sessions.

All participants will continue their usual treatment under the supervision of their referring clinical team. This involves pharmacotherapy, monitoring, and case management. They can concurrently attend any psychosocial group that does not focus on improving neurocognition.

Outcomes

Screening Measures

During screening, participants will complete the Test of Premorbid Functioning (TOPF). The TOPF is a test of premorbid intelligence estimated from reading ability and will be used to screen for intellectual impairment [35,36]. It takes

https://www.researchprotocols.org/2024/1/e52505
approximately 10 minutes to complete and is composed of a list of 70 words.

**Demographic Information**

At baseline, demographic and clinical information will be gathered, including sex, age, and date of birth. For the clinical groups, the case management team, primary and secondary diagnoses, mental health status, and list of current medications (name, dose, frequency, and route) will also be recorded.

A battery of validated assessment measures will also be delivered at baseline and follow-up for the clinical groups, as described in the subsequent sections.

**Brief Assessment of Cognition in Schizophrenia**

The Brief Assessment of Cognition in Schizophrenia (BACS) assesses 5 domains of cognition, with 6 tests taking approximately 30 minutes [37,38]. The 6 tests include list learning (verbal memory), digit sequencing (working memory), token motor task (motor speed), verbal fluency (semantic fluency and letter fluency), Tower of London (reasoning and problem solving), and symbol coding (attention and processing speed). The BACS has high test-retest reliability, is sensitive to the unique cognitive deficits associated with SSDs, and is a routine measure of change in performance over time [37].

**Schizophrenia Cognition Rating Scale**

The Schizophrenia Cognition Rating Scale (SCoRS) is a 20-item measure of cognitive difficulties in daily activities that is completed by the participant, an informant, and the interviewer at baseline and on completion of the CR program [39]. It has been found that global ratings of cognition are strongly correlated with cognitive performance, functional outcome, and functional capacity. The SCoRS has good interrater reliability [39].

**Clinical Assessment Interview for Negative Symptoms**

The Clinical Assessment Interview for Negative Symptoms (CAINS) is a measure of anhedonia, avolition, and emotional expression [40] with strong psychometric properties [41]. The CAINS has been found to have good convergent validity with the Brief Negative Symptom Scale [42]. It was developed to better align not only with the negative symptoms but also with constructs emerging from neurobiological research [43].

**Calgary Depression Scale for Schizophrenia**

The Calgary Depression Scale for Schizophrenia (CDSS) measures symptoms of depression in people with schizophrenia [44]. The scale has 8 structured questions with an additional observational item [45]. The scoring uses a 4-point Likert-type scale (0=absent, 1=mild, 2=moderate, and 3=severe), anchored by descriptors [44]. The summation of scores on all items provides a global score. The scale has good psychometric properties, identifying a major depressive episode at 82% specificity and 85% sensitivity for scores above 6.

**Brief Psychiatric Rating Scale**

The Brief Psychiatric Rating Scale (BPRS) is a semistructured assessment of 24 symptoms of schizophrenia. The 24-item anchored scale will be used. The anchored version has good psychometric properties [46]. Symptom severity is rated from 1 (not present) to 7 (extremely severe). High scores represent greater symptom severity. Based on a clinical interview, items 1-14 are based on the participants’ self-report; observed behavior is also used to rate items 7, 12, and 13. Items 15-24 are rated based on the patient’s observed behavior or speech during the interview.

**Behavioral Inhibition System and Behavioral Activation System**

The Behavioral Inhibition System and Behavioral Activation System (BIS/BAS) is a 24-item self-report measure of behavioral inhibition and activation [46,47]. A total of 13 items reflect the activation system, divided into drive, pleasure-seeking, and sensitivity to reward, and 7 items reflect the inhibition system. There are 4 filler items. Responses are rated on a 4-point Likert-type scale with a range from 1 (very true) to 4 (very false). The BIS/BAS factor structure has been validated in a large Australian community sample [48].

**Life Skills Profile**

The Life Skills Profile (LSP-16) is a measure of activities of daily living over the previous 3 months with high test-retest and interrater reliability [49]. The 16-item version was specifically developed for use in Australian public mental health services [49-51]. The items are on a 4-point anchored scale. Higher scores indicate a greater disability. A total LSP score is calculated by adding all the individual scores.

**Functional Magnetic Resonance Imaging**

Participants will complete a computerized experimental learning task with 2 conditions: a social reward condition and a nonsocial reward condition (Figure 2). The task is a simple reinforcement learning design where participants must choose from a small number of potential responses and, over repeated trials, learn about the probable consequences. In our task, participants can choose a response from 3 buttons. Each choice is associated with the unique cognitive deficits associated with SSDs but also with constructs emerging from neurobiological research [43].
During the social learning task, participants view a picture of a neutral face, with rewarded button choices changing the facial expression toward a smile and losing choices changing toward a frowning expression. Alternatively, in the nonsocial learning task, participants view a picture of a bucket half full of water. Rewarded choices further fill the bucket, while losing choices drain the bucket. The task was developed by Dr Marcus Gray for this study and implemented in MATLAB (MathWorks) using the Cogent 2000 toolbox [52]. During the functional magnetic resonance imaging (fMRI) experiment, the task is seen by participants through a tilted mirror attached to the head coil on the magnetic resonance imaging scanner. Responses are made on a commercially available, magnetic resonance-compatible response box [53]. Participants are familiarized with the task and perform 2 blocks of each condition before brain scanning.

During the fMRI, structural and functional magnetic resonance imaging images will also be acquired by a 3T Siemens Magnetom TrioTim (Siemens Healthineers) system using a 12-channel head coil. The sequences acquired and their parameters are as follows: T1-weighted imaging, T2-weighted imaging, fMRI imaging, diffusion-weighted imaging, and susceptibility-weighted imaging.

- **T1-weighted imaging**: magnetization-prepared 2 rapid acquisition gradient echoes (MP2RAGE) sequence. Time to acquire image=5:02; inversion 1=700 ms; inversion 2=2220 ms; repetition time=4000 ms; echo time=2.96 ms; voxel size=1 mm isotropic; field of vision=230 mm; and 192 slices with full brain coverage.
- **T2-weighted imaging**: fluid-attenuated inversion recovery sequence. Time to acquire image=2:44; repetition time=9000 ms; echo time=81 ms; inversion time=2500 ms; flip angle=150 degrees; voxel size = 0.72 × 0.72 × 5.2 mm; and 30 slices with full brain coverage.
- **fMRI imaging**: functional T2*-weighted blood-oxygen-level dependent images are acquired using a multiband, echo-planar sequence across the whole brain (repetition time=0.628 ms; echo time=30 ms; resolution=2.4 mm isotropic; field of view=192 mm; flip angle=52 degrees; and 54 slices with full brain coverage). During fMRI imaging, participants will complete a computerized experimental task. For each task condition, approximately 720 full brain images will be acquired, providing 1440 volumes acquired in approximately 12 minutes.
- **Diffusion-weighted imaging**: A neurite orientation dispersion and density imaging sequence with 2 shells and 90 gradient directions (B1=1000 with 30 directions and B2=2500 with 60 directions) with 6 B0 measurements will be acquired in the anterior-posterior phase encoding direction, and an additional 6 B0 measurements will be acquired in the posterior-anterior phase encoding direction. Total acquisition time=7:24; repetition time=4100 ms; echo time=75 ms; voxel size=2mm isotropic; and 68 slices with full brain coverage.
- **Susceptibility-weighted imaging**: time to acquire image=2:56; repetition time=27 ms; echo time=20 ms; flip angle=15 degrees; voxel size = 0.89 × 0.89 × 2.5 mm; and 64 slices with full brain coverage.

### Data Management

The Trial Management Group (TMG) consists of the principal investigator (FD) and associated investigators (GG and MG). All adverse events will be reviewed by the TMG and reported to the ethics committee. The process of recruitment and data management will be overseen by the TMG. The data will be securely entered and stored on the University of Queensland Data Manager repository. This trial may be subject to random auditing by the ethics committee. All protocol amendments will be reported to the ethics committee.

### Statistical Analysis

Participants’ medication dosages will be converted to olanzapine equivalents [54]. Outcome measures will be analyzed using the SPSS (version 27 or higher; IBM Corp) software package. A series of 2 (group: intervention and patient control) x 2 (time: baseline and postintervention) mixed factorial ANOVAs will be conducted to evaluate the treatment group differences for each of the outcome measures. If the normality assumption is violated, nonparametric analyses will be conducted.

For the fMRI analysis, standard preprocessing of the functionally weighted images will be carried out using the Statistical Parametric Mapping Version 12 [55]. The preprocessing steps follow: slice timing on the functional images, to correct for
differences in slice acquisition times within each volume using the middle slice as reference; realignment (estimate and reslice) on the functional images, to correct for interscan movement within each run (defined as >3 mm translation and >2 degrees rotation); coregistration of the functional and structural images; segmentation of the structural image, with heavy regularization (0.1) recommended for MP2-RAGE sequence; normalization of the resliced images into a standardized, stereotaxic space (according to the Montreal Neurological Institute template); and smoothing of normalized images with a 8 mm full-width-at-half-maximum isotropic Gaussian kernel.

The general linear model approach for event-related designs will be conducted using the Statistical Parametric Mapping Version 8 [55]. For the first-level analysis, task-related changes in the blood-oxygen-level dependent signal will be estimated at each voxel for each participant. Head motion parameters will be included as a regressor to account for participant motion during the experiment. A 1/128 Hz high-pass filter will be used to remove slow signal drifts, and a canonical hemodynamic response function with no derivatives will be selected.

Computational modeling will allow the estimation of learning parameters at each point in time, during each task condition, for each participant. We will use a variational Bayesian framework to compute how the value of each button was estimated based on the behavioral choices made and feedback received. This allows us to measure how learning occurred during the experimental task and the subprocesses that must underlie this learning. Second-level group analyses will examine how learning in patients differs from that observed in control participants and how CR therapy alters learning efficiency and the underlying neural activity. We will correct for multiple comparisons; the voxel-level threshold will be set at P<.05 family-wise error corrected.

Ethical Considerations

This trial has been approved by the Metro South Health Human Research and Ethics Committee (HREC/2021/QMS/67093). All protocol modifications and serious adverse events will be reported to the ethics committee. The study will be conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. All participation will be based on voluntary, written, and informed consent. Participants with serious mental illness can be challenging to follow up with; thus, the researchers will make multiple attempts to contact each participant at baseline and postintervention. In cases of deterioration in mental state, the intervention will be discontinued, and the participant’s treating team will be advised. Participants will be able to withdraw at any point in the study. All data used for analysis will be deidentified.

Results

This trial was registered in August 2018 and commenced recruiting in May 2022. As of September 2023, we have enrolled 11 healthy controls. We have also enrolled 16 individuals with SSDs, 15 into the CR group, and 1 into the TAU group. The projected completion of recruitment is September 2024. The projected final reporting date is September 2025. Results will be disseminated to mental health clinicians, researchers, and key stakeholders through peer-reviewed publications and presentations. In-kind funding is being provided by Metro South Addiction and Mental Health Services and the Translational Research Institute. This study received an Extraordinary Research Grant from the Princess Alexandra Hospital Research Foundation to assist with the purchase of fMRI scans in May 2023.

Discussion

Principal Findings

CR has demonstrated effectiveness in improving neurocognitive functioning in individuals with SSDs; however, the mechanisms that mediate this effect remain unclear. This study describes the unique protocol for a pre-post pilot study that aims to investigate RL as a potential mechanism for improvement following CR. In this study, we predict that at baseline, individuals with schizophrenia will exhibit deficits in RL when compared with healthy controls. Further, we predict that these deficits will be more pronounced in learning tasks that involve social versus nonsocial stimuli. This would support previous research highlighting impaired reward processing in SSDs [8], particularly in social processing [19,20]. Social reward has significant impacts on the functional outcomes of this population, and there are no current interventions that target this problem, emphasizing the importance of this area of research. Moreover, difficulty in responding to feedback has been associated with cognitive function and negative symptoms in SSDs [15-17]. Thus, we propose that by reducing RL problems, we may in turn see a reduction in negative symptoms and an improvement in cognition.

The primary focus of this study is to investigate whether RL is improved after CR in individuals with schizophrenia. Only 1 other study has specially looked at the impact of CR on reward processing in SSDs [29], and results from this research suggest CR may improve response to feedback, moderated by the severity of negative symptoms. We believe CR serves to strengthen the processes involved in rewards and motivation that are needed for participants to persevere with difficult cognitive tasks via reinforcement of learning. Learning more about the way in which CR works is important to be able to maximize the effects of the intervention. Currently, CR is undertaken for around 3 months, with 2-4 sessions per week. Understanding the mechanisms of effect may enable improvements in the programs that enable more efficient delivery of this effective intervention.

Limitations

The lack of binding of researchers and the nonrandomized assignment of participants to conditions are limitations of the study design. Treatment allocation is based on individuals’ preferences to minimize attrition. As the intervention group is self-selecting, arguably, these participants may be better functioning than the patient control group. However, we believe that this will not necessarily be the case. For instance, some higher-functioning participants might select to be in the control group due to time restraints (ie, because of full-time study or work) limiting their ability to complete the CR therapy sessions.
Nevertheless, we aim to recruit comparable participants from the patient control group and intervention group by matching for age, sex, and premorbid function across groups. There will also be some flexibility in the delivery of CR (ie, on the web vs in person and individual vs group), depending on participants’ access to the technology to run CR sessions and their preferences. This flexible mode of delivery was aimed at minimizing attrition, which has been an issue identified in the literature, in our work, and because of group therapy shutdowns during the COVID-19 pandemic [56,57].

Conclusions
It is hoped that the results of this study will contribute to the understanding of CR and RL in schizophrenia more generally. Greater knowledge of CR would seek to inform clinicians to develop more targeted interventions and, consequently, reduce negative symptoms and improve functional outcomes in individuals with SSDs. We hope to use this pilot to test the integrity of the protocol and plan for future funding, with the aim of progressing to a larger randomized controlled trial.

Data Availability
The data sets generated or analyzed during this study are not publicly available. Access to research data is restricted and governed by the Queensland Health Government.

Authors’ Contributions
FD, MC, GG, MG, VGI, VDM, and GR contributed to the design of the study. FD and VGI drafted the original protocol. FD and MC conceptualized the original idea for the study. FD and VDM will run the cognitive remediation (CR) programs. VGI and GR will conduct the cognitive assessments. MG and GR will support participants during functional magnetic resonance imaging (fMRI). No professional writers will be used. All researchers will have access to the final data set. No generative AI was used in any portion of the manuscript writing.

Conflicts of Interest
FD has received honorariums from Janssen and Lundbeck for the delivery of lectures at clinician education events.

References


52. Wellcome Centre for Human Neuroimaging. London, UK: University College London; 2023. URL: https://www.fil.ion.ucl.ac.uk/ [accessed 2023-12-21]


Abbreviations

BACS: Brief Assessment of Cognition in Schizophrenia
BIS/BAS: Behavioral Inhibition System and Behavioral Activation System
BPRS: Brief Psychiatric Rating Scale
CAINS: Clinical Assessment Interview for Negative Symptoms
CDSS: Calgary Depression Scale for Schizophrenia
CIRCUITS: Computerized Interactive Remediation of Cognition-Training for Schizophrenia
CR: cognitive remediation
fMRI: functional magnetic resonance imaging
LSPI-16: Life Skills Profile
MP2RAGE: magnetization prepared 2 rapid acquisition gradient echoes
RL: reward learning
ScoRS: Schizophrenia Cognition Rating Scale
SSD: schizophrenia spectrum disorder
TAU: treatment-as-usual
TMG: Trial Management Group
TOPF: Test of Premorbid Functioning
Protocol

Education of Patients With Atrial Fibrillation and Evaluation of the Efficacy of a Mobile Virtual Patient Environment: Protocol for a Multicenter Pseudorandomized Controlled Trial

Panagiotis Antoniou1, PhD; Eleni Dafli1, MD, PhD; George Giannakoulas2, MD, PhD; Gaukhar Igimbayeva3, MD, PhD; Olga Visternichan4, MD, PhD; Serhii Kyselov5, Prof Dr; Ivetta Lykhasenko6, MD, PhD; Dmytro Lashkul7, MD, PhD; Ilia Nadareishvili7, MD, PhD; Sergo Tabagari7, Prof Dr; Panagiotis D Bamidis1, PhD

1Lab of Medical Physics and Digital Innovation, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
21st Cardiology Department, AHEPA University General Hospital of Thessaloniki, Thessaloniki, Greece
3Orhun Medical Clinic, Almaty, Kazakhstan
4Internal Medicine Department, Karaganda Medical University, Karaganda, Kazakhstan
5Department of Internal Diseases No.1 and Simulation Medicine, Zaporizhzhia State Medical University, Zaporizhzhia, Ukraine
6Department of Propedeutics of Internal Diseases, Radiation Diagnostics and Radiation, Zaporizhzhia State Medical University, Zaporizhzhia, Ukraine
7AIETI Medical School, David Tvildiani Medical University, Tbilisi, Georgia

Corresponding Author:
Panagiotis D Bamidis, PhD
Lab of Medical Physics and Digital Innovation
School of Medicine
Aristotle University of Thessaloniki
PO Box 376
Thessaloniki, 541 24
Greece
Phone: 30 6972008122
Email: bamidis@auth.gr

Abstract

Background: Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia and is a leading cause of mortality and morbidity. Patient knowledge about AF and its management is paramount but often limited. Patients need to be appropriately informed about treatment options, medicinal adherence, and potential consequences of nonadherence, while also understanding treatment goals and expectations from it. Mobile health apps have experienced an explosion both in their availability and acceptance as “soft interventions” for patient engagement and education; however, the prolific nature of such solutions revealed a gap in the evidence base regarding their efficacy and impact. Virtual patients (VPs), interactive computer simulations, have been used as learning activities in modern health care education. VPs demonstrably improved cognitive and behavioral skills; hence, they have been effectively implemented across undergraduate and postgraduate curricula. However, their application in patient education has been rather limited so far.

Objective: This work aims to implement and evaluate the efficacy of a mobile-deployed VP regimen for the education and engagement of patients with AF on crucial topics regarding their condition. A mobile VP app is being developed with the goal of each VP being a simple scenario with a set goal and very specific messages and will be subsequently attempted and evaluated.

Methods: A mobile VP player app is being developed so as to be used for the design of 3 educational scenarios for AF management. A pseudorandomized controlled trial for the efficacy of VPs is planned to be executed at 3 sites in Greece, Ukraine, and Kazakhstan for patients with AF. The Welch t test will be used to demonstrate the performance of patients’ evaluation of the VP experience.

Results: Our study is at the development stage. A preliminary study regarding the system’s development and feasibility was initiated in December 2022. The results of our study are expected to be available in 2024 or when the needed sample size is achieved.

Conclusions: This study aims to evaluate and demonstrate the first significant evidence for the value of VP resources in outreach and training endeavors for empowering and patients with AF and fostering healthy habits among them.
atrial fibrillation; virtual patient; scenario based learning; technology enhanced learning; mHealth; mobile health; patient engagement; patient education; cardiac arrhythmia; mortality; mobile application; mobile app; health education; randomized control trial; cardiology; cardiac; heart; Greece; Ukraine; Kazakhstan; clinical decision support systems; CDSS; virtual patient scenario; myocardial infarction; arrhythmia; stroke

Introduction

Impact of Atrial Fibrillation and Non–Vitamin K Oral Anticoagulants and the Role of Integrated Management

Atrial fibrillation (AF) is the most frequent long-term cardiac arrhythmia in adults, presenting a considerable burden on patients, health care systems, societal health, and the global health economy [1]. Because AF is associated with significant morbidity and mortality, numerous research efforts and resources are being directed toward gaining more detailed information about the mechanisms underlying AF, its natural course, and effective treatments, and new evidence is being generated and published on a regular basis [2].

The current estimated prevalence of AF is between 2% and 4% [3], with a 2.3-fold increase projected due to the general population’s increased longevity and the intensive search for AF [4]. AF is a well-known risk factor for thrombus development in the left atrium and eventual embolism on the left side. AF raises the risk of stroke by 5 folds, but the risk varies depending on the presence of stroke risk factors or modifiers [2]. Oral anticoagulants (OAC) are the cornerstone of AF treatment. Vitamin K antagonist (VKA) medication (mainly warfarin) lowers the risk of stroke and death by 64% and 26%, respectively, when compared to control or placebo [5]. Non–vitamin K oral OACs (NOACs) have outperformed VKAs in most therapeutic circumstances. NOAC medications do not have the practical constraints of VKAs, such as a small therapeutic window, interactions with food and other treatments, and the need to monitor coagulation levels. In 4 large randomized controlled trials (RCTs) with patients with AF, NOACs were compared to warfarin [6-9]. They were demonstrated to be at least noninferior to VKA therapy for the prevention of stroke or recurrent venous thromboembolism and were associated with a lower risk of bleeding. In a meta-analysis of these RCTs, NOACs were associated with a 19% significant reduction in the risk of stroke or systemic embolism, a 51% reduction in the risk of hemorrhagic stroke, and a similar reduction in the risk of ischemic stroke compared to VKAs. NOACs are also associated with a 10% reduction in all-cause mortality [10].

To provide optimal medical treatment to patients with AF, integrated management necessitates a patient-individualized care route. Treatment choices should be reviewed and a management plan agreed upon with health care experts in this patient-centered approach [11]. Treatment is liable to alter over time as new symptoms, risk factors, disease progression, and novel medicines emerge. To prevent stroke and improve symptoms, it is critical to consider optimizing resource usage. The initial stage in shared decision-making should be to investigate the patients’ preferences [12]. The importance of stroke prevention and rhythm control among patients, as well as the corresponding risks of death, stroke, and significant bleeding, should be properly appraised for shared decision-making.

Patients’ awareness of AF and its management is sometimes restricted, especially when first diagnosed, because the majority of treatment decisions must be addressed and made. Furthermore, controlling patients’ expectations of treatment goals, as well as ensuring that patients are correctly informed about treatment options, how to better adhere to therapy, and potential repercussions of nonadherence, are critical to increase adherence [2]. Thorough education of physicians on these approaches, as well as correct adherence and active engagement of the patient in the treatment process, are critical to the success of each treatment plan. Soft health interventions for education and empowerment of both clinicians and patients are critical in this environment.

Education and Other Soft Interventions Involving Mobile Health Apps in Cardiology

There has been an explosion of mobile health (mHealth) apps during the previous decade, with an estimated 3.7 billion downloads globally between 2013 and 2017 [13], many of which are intended for AF. A 2020 systematic review revealed around 11,152 articles related to mHealth apps for AF but only included 9 studies with real outcomes about mHealth therapies for AF in its results synthesis [14]. This indicates the prevalence of such solutions, as well as the dearth of data on their efficacy and impact. Only a few apps have been evaluated formally [15-17]. It should be emphasized that there are few apps that are specifically intended for patients. There are numerous informative apps that patients can use; however, a preferred option would be the development of a specific app for the treatment of a given ailment in order to help afflicted individuals in a more appropriate way [18].

Clinical decision support systems that digitize and provide evidence-based recommendations, therapeutic pathways, and algorithms to facilitate individualized, timely, and evidence-based treatment could be a valuable aid in the holistic management of AF. To improve patient-integrated AF management [19], the MobiGuide project [20] and numerous applications have been deployed. The European Society of Cardiology—Characterizing Atrial fibrillation by Translating its Causes into Health Modifiers in the Elderly Consortium partnership offers a smartphone or tablet app for patients with
AF; however, it has yet to be tested prospectively [21]. Contradictory findings highlight the need for more properly designed trials, including evaluation of the intervention's influence on clinical outcomes [22].

As a result, the scope and impact of mHealth apps for AF, as well as the level of patient and health care professional (HCP) engagement and acceptability, are currently unknown. HCP engagement refers to information sharing between the patient and provider, shared responsibilities in decision-making processes, and support of patients' choices and acceptability to a degree to which an intervention is approved by most HCPs. Given both patients and HCPs may easily use these applications, it is critical to understand their scope and content and their acceptability among users, and to investigate the purpose and results of app adoption and usage.

**Virtual Patients in Health Care Education**

Virtual patients (VPs)—interactive computer-based clinical scenarios for the purpose of medical training, education, or assessment—have been increasingly used as learning activities in current health care education, particularly in teaching decision-making through scenarios [23]. Because VPs can enhance cognitive and behavioral skills, they have been successfully incorporated in undergraduate and postgraduate curricula [24]. With the rising use of VPs, there are opportunities for pedagogical synergies to allow trainees of diverse categories to practice in realistic and safe learning contexts [25]. The key characteristics of VP systems are that they enable repetitive and intentional practice by any student, regardless of time or physical location, and that mistakes are not fatal. These opportunities provided by VPs in current medical education, combined with positive evaluation results from various studies demonstrating that they may improve cognitive and behavioral skills better than traditional methods [24], have resulted in a widespread trend toward VP creation and use among academic institutions [26]. Furthermore, VPs enable the production and usage of more game-based educational content, which provides the student more exploratory flexibility and provides a different area for case-based content in current medical education.

The widespread adoption of these digital technologies, not only in medical student education but also in the health care community in general and in the patient community, has undoubtedly been limited thus far, but it has the potential to educate and psychologically support high-risk patients and vulnerable populations in these new and unprecedented circumstances. Because numerous academic institutions have the VP resources and expertise in their execution, such effective educational content might be simply repurposed in a more patient-centered format and be used more widely by the patient community as an educational aid.

**Study Aim and Objectives**

In this technological setting, we propose the development of a holistic approach to patient engagement and education based on the fundamentals of AF. The DEEP-RAFT (Doctors’ Education, Empowerment of Patients, Regarding Atrial Fibrillation and Venous Thromboembolism) project would generate a suite of educational and informational interventions along an axis created by 2 poles: digital content creation and evidence-based educational impact. This effort will be based on a suite of digital teaching resources, in the form of VPs, co-designed by continuous and immediate involvement of health care specialists, health care policy influencers, and patients. This strategy seeks, first and foremost, to develop materials that are more patient-centered and address realistic problems relevant to the health care systems involved in the project. The focus of this effort is on the second topic, evidence-based educational impact, which attempts to demonstrate the educational efficacy and acceptance of the generated resources as they may be used in a diverse but targeted set of education and outreach activities.

In practice, a mobile app for natively deploying VPs on mobile devices is developed, along with a list of relevant VPs. In the following parts, we will provide these concrete and intangible tools and approaches to contextualize the protocol that is described below.

**Methods**

**Study Design**

This is a 2-arm, parallel-group pseudorandomized controlled trial that will be performed at 3 sites—Greece, Ukraine, and Kazakhstan—and will include the evaluation of the educational VP interventions. This will be conducted in 2 axes that correspond to the primary outcomes of the study. The first axis aims to determine the efficacy of the educational VP interventions, while the other will involve assessing the acceptance and opinions of the participants about the VP modality for education and information purposes. Correspondingly, the primary outcomes of the study include the efficacy and the acceptance of the VP interventions, and the secondary outcomes include the opinions of the participants about the VP modality for education and information purposes. The aims of the study are summarized in the following research question: are patients with AF better educated about their post–acute phase lifestyle changes and needs by using mobile virtual scenarios compared to conventional patient education methods?

A preplot arm of the study will be initially conducted in Greece and will consist of hospitalized patients due to AF episodes, all of which should complete a short quiz. From among the patients who complete the quiz, half of the patients will be randomly allocated to the control (normal clinical information) cohort, while the other half will be informed with the help of VP vignettes.

After the preplot arm, the pseudorandomized controlled trial will be conducted multicentrically. The same process, as in the preplot arm, will be followed for recruiting the core sample, exploring the efficacy and impact of VPs in educating patients with AF. Since this trial will have been conducted during the COVID-19 pandemic, all national and international health and safety protocols will be followed.
Inclusion and Exclusion Criteria and the Recruitment Process

The inclusion and exclusion criteria are outlined in Textbox 1. The dropout criterion is withdrawing consent during the study period.

In the Greek pilot where randomization of patient cohorts would occur, a simple coin toss algorithm will be used, implemented by Excel’s (Microsoft Corp) RANDBETWEEN function. This will allocate patients between the VP education cohort and the control (normal clinical information) cohort. In the Ukraine pilot, a sampling of convenience will allocate most of the participants to the VP cohorts and fewer to the control (normal clinical information) cohort. In the Kazakhstan pilot, all accepted patients will be allocated to the VP intervention cohort. While this decision by the medical team of this center makes impossible the conduct of a distinct local evaluation of impact, we chose to include the sample in the multicentric data processing part of the study. It should be noted that evaluation results will be extracted on a per-site basis only in the Greece and Ukraine arms of the study. In the multisite comparison, the whole sample of intervention patients (including the totality of the Kazakh cohort) will be compared to the totality of the sample of control patients from Greece and Ukraine cohorts.

Patients who meet the requirements for that study will be informed about the study and will be asked if they would like to participate in it. They will then sign the consent form to take part in the study. Patients who may withdraw their consent during the study will be excluded from the analysis.

Textbox 1. Study inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Patients with a history of paroxysmal, persistent, or permanent atrial fibrillation</td>
</tr>
<tr>
<td>• Access to an internet connection and adequate equipment</td>
</tr>
<tr>
<td>• Mastery of the country’s first language</td>
</tr>
<tr>
<td>• Informed consent provided by the participant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Organic or symptomatic mental disorders</td>
</tr>
<tr>
<td>• Alzheimer disease</td>
</tr>
<tr>
<td>• Mental and behavioral disorders due to alcohol abuse</td>
</tr>
<tr>
<td>• Mental and behavioral disorders due to drug abuse</td>
</tr>
<tr>
<td>• Refusal of patients to receive basic drug therapy</td>
</tr>
</tbody>
</table>

Intervention Cohorts

The full pilot arm of the study will be conducted multinationally in Greece, Ukraine, and Kazakhstan (Figure 1). In Greece, a pool of 700 patients will be reached by phone, and they will participate in the trial from their home. From among the patients contacted remotely, and from those who respond, those who would be eligible will be asked to complete the intervention. From among the patients who choose to complete the quiz, half of them will be randomly allocated to the control (normal clinical information) cohort, while the other half will be informed with the help of VP vignettes. In Ukraine, 160 patients will be reached and will be divided into the control and intervention groups. Finally, at the Kazakhstan center, 190 patients will be contacted.

Given that all recruitment and adherence criteria are followed, sufficient sample sizes for control and intervention groups will be ensured so as to use the Welch t test for analysis, which is robust to large sample size inequalities, in order to not exclude any useful data from the generalized multicenter VP cohort.
Blinding (Masking)

Participants and trial personnel will not be blinded after the point of assignment to interventions because of the nature of the interventions and outcomes assessed. Participants will know which group they will belong to because the group-specific intervention will follow immediately after randomization.

The Intervention

The 2 groups that will be compared are as follows.

The VP Cohort

In this cohort, doctors of the research team will contact patients and ask them to participate in the VP vignette case (see The VPs section for a detailed description of the virtual scenarios used in the intervention, and the The Technical Architecture of the Mobile App section for details of the mobile app that will be used for delivering the intervention). At the end of the case questions, further information will be provided.

The Control (Normal Clinical Information) Cohort

In this cohort, doctors of the research team will contact patients and provide the protocol—which will include dictated information for patients with AF.

The Technical Architecture of the Mobile App

A mobile VP player app has been created, which is capable of tracking the following user data: a detailed log of pathway detection and tracking, time spent in each node of the VPs or mobile VP, tracking of milestones, rate of successful completion of the VPs or mobile VP, and connection to learning outcomes. The player follows a flat development approach and should be available as a progressive web application. A progressive web application is a type of web-based application software that is built using standard web technologies such as HTML, CSS, and JavaScript. It is aimed to be compatible with any platform that supports a standards-compliant browser. The ability to work offline, receive push notifications, and access device hardware will enable the creation of user experiences similar to those found in native applications on desktop and mobile devices. Because a progressive web application is a subset of a web page or website known to be a web application, neither developers nor users should be required to install web applications through digital distribution systems such as the Apple App Store or Google Play.

Technically, communication between the backend and the front end player is accomplished through the use of the JSON application programming interface standard [27], which is an...
extremely efficient method of exchanging data over slow networks (e.g., mobile phone networks).

The system’s appearance and feel are based on Bootstrap, a free and open-source CSS framework for developing responsive, mobile-first front end web applications. It would be optimized for smartphone screens so as to provide a “mobile-first” user experience.

**The VPs**

Three scenarios will be explored by the participating medical teams. All will be focused on the patients, since these are the target group of the VPs. A tabulated outline of these scenarios is presented in Table 1. The first scenario should be dealing with a chronic case of AF, including guidance for good medication adherence, systematic medical appointments, and correct communication. The second scenario would be about detecting early onset and prevention of paroxysmic AF. The third scenario will deal with a case of AF that involves lifestyle-compounding factors to the disease (smoking, drinking, and unhealthy eating), as well as more heavy complications that are significantly more probable in such cases.

<table>
<thead>
<tr>
<th>Scenario theme</th>
<th>Educational objectives (after encountering this educational virtual scenario vignette, the patient will be able to…)</th>
</tr>
</thead>
</table>
| Chronic case of AF                               | • Identify correct medication adherence practices  
• Identify the correct frequency regimen of doctors’ appointments  
• Identify correct dietary restrictions                                                                 |
| Early-onset and paroxysmic AF                    | • Identify initial onset of symptoms of AF  
• Identify timely medical consultation practices  
• Recognize the correct medicinal adherence procedure for the condition  
• Identify the correct exercise intensity to manage the condition |
| AF with lifestyle-compounding factors to the disease (smoking, drinking, and unhealthy eating) | • All educational objectives of “Chronic case of AF”  
• Identify risks of high-impact complications  
• Identify best practices for recovering from missed doses of medical treatment |

*AF: atrial fibrillation.

Iterative brainstorming between the medical experts of all study centers will produce the detailed VP scenarios in the mobile digital platform (for details, refer to the previous section, The Technical Architecture of the Mobile App). After an internal review of these scenarios by the multinational expert panel, a selection would be made for the final case to be used. In that context, the team will choose to simply use 1 scenario in the multicenter multinational cohort trials. The most appropriate scenario would have the following characteristics: (1) it should be the most relevant—a significantly larger proportion of patients should be targeted; (2) it should be the most clinically impactful—it is important and useful for patients with AF to be aware of AF complications; and (3) it should be the most educationally important—both patients who fall within the parameters of the described case and those who do not would benefit from information and preventative knowledge of the impact of AF.

An acceptance rate of 75% by medical experts for each criterion for each scenario should be reached to be eligible for recruitment. Moreover, this patient education intervention will be based on techniques from the behavior change technique taxonomy—an international consensus for the reporting of behavior change interventions [28].

After this selection, a process of localization and adaptation to the specifics of each center (Greece, Ukraine, and Kazakhstan) would be conducted.

**Evaluation Design**

**Evaluation Instruments**

For educational efficacy, a multiple-choice questionnaire (MCQ) for knowledge retention will be used (Textbox 2). The choice of questions was based on the European Society of Cardiology’s Guidelines for Management of Atrial Fibrillation [29]. This instrument will be translated for all participating centers in Greece, Ukraine, and Kazakhstan. The translation and evaluation of the translation will be performed by teams of bilingual experts. The instrument in its original version will be provided to bilingual persons in alternating language order and will be assessed accordingly. Scoring each questionnaire follows a simple process of allocating a numerical score equal to the number of choices in the MCQ to weigh each response for randomly selecting the correct question. For example, a question that has 4 possible responses in the MCQ will be scored 4 points if answered correctly, while a question that has 5 possible answers will be scored with 5 points if correctly answered.
1. **Atrial fibrillation:**
   - Is the most common cardiac arrhythmia
   - More often concerns older people, but can occur in any age
   - May be related with thyroid disorders
   - All of the above*

2. **The most common symptom of atrial fibrillation is:**
   - Chest pain
   - Palpitations*
   - Dizziness
   - Blurry vision

3. **How can atrial fibrillation be diagnosed?**
   - Following an electrocardiogram evaluated by a cardiologist*
   - By describing symptoms of the arrhythmia to the doctor
   - By checking the indication ‘arrhythmia’ of the blood pressure device
   - All of the above

4. **What is the most important complication a patient with atrial fibrillation not receiving treatment may suffer?**
   - Fainting spells
   - Myocardial infarction
   - Lethal arrhythmia
   - Ischemic stroke*

5. **Treatment with oral anticoagulants always mandates frequent blood tests:**
   - True
   - False*

6. **The patient with atrial fibrillation that visits a cardiologist:**
   - Probably does not need oral anticoagulation to prevent stroke
   - Always needs treatment to cure the arrhythmia
   - Cardiologist? There is no need to see a doctor
   - Might have to be admitted to the hospital at the time of Atrial Fibrillation diagnosis *

7. **When the patient with atrial fibrillation is being treated with an oral anticoagulant:**
   - It is better to receive the reduced dose so as to avoid bleeding
   - This is always stopped at 3 months, since the danger for stroke is gradually reduced
   - It is fine if he/she occasionally misses a dose
   - He/she has to adherently receive the right dose of the drug, as prescribed by his/her treating physician*

8. **The newer oral anticoagulants:**
   - Are at least as safe and effective as warfarin in preventing stroke
   - Have to be taken every day on a fixed schedule, so as to be effective
   - Do not have important interactions with other drugs or food, in contrast to warfarin
   - All of the above*

9. **Oral anticoagulants:**
   - Are not necessary, in case the patient already receives other blood thinning medication, such as aspirin
Are the most effective treatment in preventing stroke*
Do not have any significant side effect
All of the above

10. A patient that is under warfarin:
- Cannot switch to a newer oral anticoagulant if he/she has good anticoagulation control (international normalized ratio within the desired range)
- Can follow an unrestricted diet
- Can get advice from his doctor regarding treatment with a newer anticoagulant, so as there is no need for frequent blood tests*
- Always has good anticoagulation control (INR within the desired range) if he/she receives his medication in a fixed dose

Project Management
All project members will meet remotely every week to work through advances and challenges together and to provide methodological support to remain aligned with the protocol. Researchers will be hired and trained, regulate safety conditions, and oversee the data collection and analysis. The researchers will prepare the data collection tools and perform data collection, and ensure that the materials required are adequate and functional. The senior Greek PI (PDB) will coordinate the overall project.

Ethical Considerations
Ethics approval has been obtained from the Bioethics committee of the Medical School of the Aristotle University of Thessaloniki (178/19-3-2020). This study will be conducted in line with the tenets of the Declaration of Helsinki, and no participant will be randomized unless written informed consent is available for that participant. Participants can withdraw from the trial at any time and will be informed and assured of such right. This study follows the principles of data protection and management described in the European Union’s General Data Protection Regulation.

Confidentiality
All personal and collected data will be treated as confidential at all stages of the study and will be stored separately. The electronic data will be saved with metadata in university network drives, which are protected by usernames and passwords. The participant ID list that links the study participants and research data will be disposed of after 15 years. Institutions hold the ownership of registry data.

Results
The trials will start in 2024 and are expected to end later that year or in early 2025 or when the needed sample size is achieved. The initial results are expected by March 2024.

To assess the results from the pretrial evaluation questionnaires, the Welch t test will be conducted. Of note, we decided to conduct the Welch t test because some of our sample sizes will be heavily unequal between intervention and control groups. The standard Student t test is robust to inhomogeneity of variance when sample sizes between cohorts are the same; however, this is not true for largely differing cohort sample sizes. The Welch t test does not assume homogeneity of variance and, hence, is robust to widely varying sample sizes [30,31].

Discussion
Expected Findings
The results of this study aim to demonstrate the efficacy of the VP educational modality in transferring knowledge in an impactful way so that it would be useful and retained by the learner. The rationale for this kind of expected efficacy may be attributed to various factors. The impact of information passed through narrative vehicles has been identified early on [32]. Additionally, the initiative that the learner has to guide the narrative through their choices facilitates engagement through 2 avenues: one avenue is the agency that the learner has over their narrative exploration, and the second one, dependent on the first but not identical, is the ability to direct the narrative toward educational needs that the learner may have. These factors are compounded by the immediacy and ease of access that the mobile delivery platform offers, which can create an engaging and user-friendly experience that may amplify the educational impact.

This work will focus on knowledge retention and efficacy and not on perceived changes in quality of life. The impact of AF itself in the quality of life of patients is well documented with a validated questionnaire that has been available since the last decade [33]. A cursory search with the keyword “AFEQT” revealed more than 550 references on Google Scholar, including several reviews (for characteristic examples, see Kotecha et al [34] and Parameswaran et al [35]).

On the other hand, there is a significant body of literature that has identified the perceived impact of impactful patient education in their risks for serious complications and quality of life. AF, when first diagnosed, is an overwhelming situation for the patient, and reliable information is one of the first requirements for alleviating the initial possible shock [36,37]. Furthermore, lack of knowledge leads patients with AF to have significantly skewed perceptions about the importance of their condition and the true risks that stem from this potentially life-threatening, possibly chronic, condition.

Multiple studies have demonstrated that patients do not identify the possibility of stroke as an acute complication of AF; a lot of them cannot even identify their arrhythmias as AF, and they
do not recognize AF as life-threatening even though they receive verbal or printed information about their condition [38,39]. In that context, it is very important to constantly explore information and educational avenues that are impactful and engaging for patients who need reliable and immediately absorbable information. This study aims to demonstrate that in this very important aspect of knowledge retention, the approach of narrative virtual scenarios is one that could provide a distinct impact advantage over the conventional paper-based or verbal information to the patients.

Limitations

The study’s core limitation lies with the knowledge retention questionnaire. This is not a validated instrument. While we were aware of a formal validated instrument (the Jessa Atrial fibrillation Knowledge Questionnaire [40]), the knowledge spectrum that it covers is far wider than what our VP vignette may cover. Given that verbal and printed conventional information covers all relevant material, while our VP may cover specific critical topics related to AF, using a wide instrument such as the Jessa Atrial fibrillation Knowledge Questionnaire would run the risk to evaluate knowledge retention gaps that the VP vignette cannot cover. While a counterargument can be made that we are thus narrowing the evaluation scope to the strong points of our VP resource, the concise focus of the VP vignette is itself an argument for implementing it as a more effective educational tool in personalized and focused endeavors for patient empowerment. As a follow through too, we aim to address the second weakness of our study, which is the lack of a qualitative exploration about the acceptance of patients regarding the electronic medium of mobile devices and the modality of VPs in comparison to other existing modalities such as video demonstration or even gamified virtual environments.

Even given these limitations, however, this study can provide evidence for the comparatively better efficacy of the VP modality in mobile media for impactful and effective information and for empowerment of patients with AF. This study would be the first evidence-based step to initiate this process toward better informing and subsequent empowerment of patients with regard to the management of their disease.

Conclusions

This project can generate new knowledge and relevant results for a deployed VP regimen for the education and engagement of patients with AF on crucial topics regarding their condition. A 3-center pseudorandomized controlled trial could add data to the evidence regarding the effects of interventions using VP resources in outreach and training endeavors for empowering patients with AF.

Acknowledgments

This work is partially supported by Pfizer’s independent grants for learning and change on Improving the Identification and Management of Patients with Atrial Fibrillation in Order to Reduce Stroke Risk and Venous Thromboembolism in Eastern Europe, project DEEP-RAFT (grant ID 42537945), and the Training and Ubiquitous Outreach, Using Social INnovation and Co-Design of Learning resoUrces and Digital Engagement, project TOUS-INCLUDE (grant ID 63829625).

Conflicts of Interest

GG received fees for lecture travels or research from Bayer, Boehringer Ingelheim, Pfizer, and LeoPharma. The other authors declare no conflicts of interest.

References


27. JSON API. URL: https://jsonapi.org [accessed 2020-11-14]


Abbreviations

AF: atrial fibrillation
DEEP-RAFT: Doctors’ Education, Empowerment of Patients, Regarding Atrial Fibrillation and Venous Thromboembolism
HCP: health care professional
MCQ: multiple-choice questionnaire
mHealth: mobile health
NOAC: non–vitamin K oral coagulant
OAC: oral anticoagulant
RCT: randomized controlled trial
VKA: vitamin K antagonist
VP: virtual patient
Development and Testing of an Electronic Diabetes Diary Integrated With a Hospital Information System for Individuals With Type 2 Diabetes Mellitus: Protocol for a Mixed Methods Study

Geena Skaria¹*, MSc; Bhageerathy Reshmi¹*, MSc, PhD; Sabu K M¹*, MSc, PhD; Sahana Shetty²*, MD, DM; Vani Lakshmi R³*, MSc, PhD

¹Department of Health Information Management, Manipal College of Health Professions, Manipal Academy of Higher Education, Manipal, India
²Department of Endocrinology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, India
³Department of Data Science, Prasanna School of Public Health, Manipal Academy of Higher Education, Manipal, India
*all authors contributed equally

Corresponding Author:
Bhageerathy Reshmi, MSc, PhD
Department of Health Information Management
Manipal College of Health Professions
Manipal Academy of Higher Education
Manipal, 576104
India
Phone: 91 9880530701
Email: reshmi.b@manipal.edu

Abstract

Background: Type 2 diabetes mellitus (T2DM) is one of the leading noncommunicable diseases that require diabetes self-management (DSM) practices. This study proposes to develop a customized mobile health (mHealth) app integrated with a hospital information system (HIS) to enable real-time, two-way transfer of information between the patient and physician. The captured information in the electronic health record will facilitate physicians to have a chronological account of the patient’s diabetes history and enable tweaking of the treatment.

Objective: The objectives of the study are (1) to develop the HIS-integrated Electronic Diabetes Diary (EDDy) per the end-user expectations at a tertiary care hospital in a south Indian state with a high prevalence of T2DM and (2) to evaluate and test adherence to EDDy in the management of T2DM.

Methods: The study will be carried out in 3 phases. Phase 1 involved in-depth interviews with primary end users to gather information regarding their expectations from the hospital-based EDDy. Phase 2 will use this information to develop a customized mHealth app using an iterative model of software development. Phase 3 will involve a pre- and posttest design: the developed app will be tested among consenting patients, where physicians will receive the patients’ data through the HIS-integrated mHealth app. The pre- and posttest values will be analyzed for adherence leading to improvement in patients’ self-management of blood glucose, user experience, glycemic control, and clinical utility.

Results: Phase 1 was completed on November 28, 2023. Phase 2 commenced in December 2023 and will end in May 2025. Phase 3 will follow afterward.

Conclusions: The proposed app will include a convenient and simple alert system that enables the patient to test glucose values at self-selected intervals, provide grading options to enter diabetic-related complications, enhance patients’ knowledge of tracking and managing the complications of diabetes, and help in maintaining the visual representation of glucose values and complications. The simplicity and usability of the modules are its novelty, which may motivate the patients to keep track of their glucose values and help them attain better health outcomes.

Trial Registration: Clinical Trial Registry India CTRI/2023/03/051077; http://tinyurl.com/4tau4ndb
International Registered Report Identifier (IRRID): PRR1-10.2196/50732

(JMIR Res Protoc 2024;13:e50732) doi:10.2196/50732
KEYWORDS
self-management of blood glucose; SMBG; diabetes self-management; DSM; personal health records; electronic diabetes diary; glycemic control; patient adherence; digital health

Introduction
In India, 11.8% of the population have diabetes per the National Diabetes and Diabetic Retinopathy Survey released by the Ministry of Health and Family Welfare in 2019, and this is projected to have a steady rise of 3.8% annually [1]. This prevalence is higher in the southern states and union territories of India [2]. The Indian Glucose Monitoring Device Market Reports of 2020 [3] suggested that the use of glucometer was always on the rise, and it was specifically noticed that there was an exponential increase in its sale during the COVID-19 pandemic period. Technology-enabled self-management has been increasingly recommended, along with proper medication and lifestyle modification, to tackle diabetes. There are numerous diabetes management apps available in the Google Play store; however, subscription to such apps is not proportional to the population with diabetes [4]. Literature from other parts of the world suggest that the adoption of digital diabetes-monitoring solutions is expected to reach 21% by 2027 [5].

Diabetes is an irreversible chronic condition that need to be well managed by altering lifestyles, food habits, and physical activities [6]. However, a standard diabetes management regimen may not be feasible to everyone around the globe, as each of these factors varies according to geographical locations and general living conditions [5]. Using customized information that is specific to the local lifestyle may be more effective and easier to ensure adherence. Existing literature suggests that nonadherence to treatment regimen and irregularity in following a diabetes-friendly lifestyle leads to poorly managed diabetes and poor disease outcomes [7]. To enhance the existing measures to manage T2DM, efficient methods needs to be formulated to ensure adherence to treatment regimen [8]. The success of such activities requires active involvement of the patients and caregivers. They should be motivated to follow T2DM management instructions from health care professionals.

This study setting already has an existing system of manual diabetes information documentation that expects its patients to record their self-management of blood glucose (SMBG) practices systematically and submit them to the treating physician during consultation. These glucose monitoring records are used by the physician to evaluate the patient’s glycemic status and make appropriate clinical decisions. This study proposes to automate this process and make it more efficient for the patients and physicians. The proposed electronic diary is expected to reduce the loss of documented values, minimize errors, as well as motivate the patients to effectively make use of technology and manage their condition.

A pilot study was conducted to analyze the knowledge and attitude to digital diabetes management practices among 50 existing patients with diabetes, which indicated that ignorance and technological illiteracy are the main reasons behind nonuse of digital apps among them. Those who used digital apps, such as mobile health (mHealth) apps, suggested that the information available is more general in nature and that they find it difficult to navigate and comprehend the information, eventually leading to the discontinuation of its use.

Methods
Overview
The study will be carried out in the outpatient department (OPD) of the endocrinology department at a tertiary care hospital in a south Indian state, and it is divided into 3 phases. In phase 1, a qualitative study was conducted; in phase 2, the mobile app with be developed and integrated with a health information system (HIS); and finally, in phase 3, the developed app will be tested for its feasibility using a pre- and posttest design.

The study is based on the principles of the “theory of change,” where stakeholders will be consulted through in-depth interviews (IDIs) to identify the deterrents to acceptance and adherence of electronic diabetes-monitoring practices, potential challenges, as well as ethical issues related to electronic information sharing. This feedback will provide the cues for the development of a customized electronic diabetes-monitoring system. The study would also attempt to identify the reasons that would lead to the discontinuation of electronic monitoring practices among patient with T2DM and address them by developing a more user-friendly mHealth app. This study will also analyze the patient acceptance of the mHealth app and their adherence to SMBG practices.

The model will then be developed and integrated with an HIS. Finally, it will be tested among patients with T2DM regarding their acceptance and adherence to using the Electronic Diabetes Diary (EDDy), its usability, contextual relevance according to user expectation, and clinical utility in terms of glycemic control.

Ethical Considerations
This study has been approved by the Institutional Ethics Committee (IEC; IEC No:223/2022). The participants personal information will be anonymized, and all privacy, confidentiality, and compensation policies will be adhered to as per the IEC guidelines.

Pilot Study
As a precursor to this study, a pilot study was conducted in a tertiary care hospital among 50 patients with T2DM to understand the existing personal record-keeping practices, the challenges they face, and their awareness of personal record keeping. The perception, expectations, and concerns of the target population were studied using a validated questionnaire survey. Patients with T2DM who are within the age group of 40-65 years were included in the study. The study setting encouraged patients with T2DM to document their SMBG values by providing a data sheet to enter the values and asking patients to bring it with them when they come for a doctor’s appointment. It is a valuable tool for the doctor to assess the
values and make better clinical decisions in the management of the disease.

A cross-sectional study was conducted, and a purposive sampling technique was followed to select the participants. A mixed-type questionnaire (both open- and close-ended questions) was used to collect data.

**Phase 1**

**Study Design**

IDIs with key stakeholders was carried out to identify the reasons for nonadherence, expectations from EDDy, and mandatory requirements to develop the intervention mechanism. The key stakeholders included the following participants:

- Patients with T2DM
- Their caregivers
- Treating doctors
- Diabetes counselors (includes nursing counselors, diet counsellors, and physiotherapists)

**Sample Selection**

The IDI participants were enrolled according to the following criteria: patients who own an Android smartphone and those who regularly attend doctor’s appointments (i.e., those who visit the OPD at least once in the last 6 months). Patients who are illiterate, can understand the local language (Kannada or Tulu) or English, aged 40-65 years, and have hemoglobin A1c values between 7% and 10% were selected to ensure homogeneity. Patients with gestational diabetes and other terminal illnesses were excluded. The IDIs were conducted among the immediate caregivers of the patients (i.e., those who are actively involved in the care of the patient). To understand the practical aspects of the treatment and management of patients, physicians, nurses, and counsellors were also interviewed.

Even with the availability of various diabetes management apps, studies suggest they are not very popular among the general population. The reason for this underuse and nonadherence, along with their expectations from mHealth apps, are the key areas that were addressed among patients with T2DM and their caregivers. The doctors and diabetes counselors were interviewed to understand the challenges they face while managing patients with T2DM and their SMBG practices.

**IDI Process**

The IDIs focused on the perspectives of patients with T2DM and their caregivers along with physicians and diabetes counselors who are actively involved in the care of patient with T2DM. Additionally, IT and health information management professionals were also consulted on ethical challenges and technical expertise. The IDI guide was designed based on the objectives of the study and the content was validated by experts. The IDI process is shown in Figure 1.

Qualitative data generated during the IDIs will be analyzed by using ATLAS.ti software (version 8; ATLAS.ti), and a report will be prepared. Based on the report, the researchers and the subject experts will identify the most feasible options that will be considered for incorporation into EDDy.

**Phase 2**

**Development of EDDy**

Phase 2 is the main part of this research, where the software will be developed based on inputs from the IDIs and the literature. The mobile app will be developed on Android Studio (Google) using an iterative model design.

The iterative model (Figure 2) is one of the Software Development Life Cycle models, where the development of a system goes through repeated cycles (iterative) and is conducted in smaller portions at a time (incremental) [9]. As and when the different modules of the app are developed, they will be tested and reviewed against the end user’s expectation, and accordingly, modifications and further development will be done. This process will enable easier debugging and fixing of errors as and when tested.
The development and testing of the app will involve development, pilot-testing, and debugging at each phase. Each iteration will be subjected to the development, testing implementation, and analysis of each component of EDDy. The components for incorporation into the software will be validated by the following experts:

- An endocrinologist
- Qualitative experts
- The IT team at the hospital

The expected pattern and time period for phase 2 are as follows.

**System Development**

**Step 1: Requirements**
This is the phase where the stakeholders’ requirements are enquired and compiled, which was primarily covered in phase 1. The captured data will be systematically entered into Microsoft Excel for further evaluation.

**Step 2: Analysis**
The compiled information will be divided into feasible and nonfeasible items as per the system development and HIS-interfacing requirements. This step will be conducted with the researchers and technical experts.

**Step 3: Designing and Development**
This step will be done by a software developer with the support of a researcher. Once the feasible requirements are sorted and finalized by the researcher, the design of the app will take place. First, a prototype of the system will be developed, and the requirements for each module of the app will be determined. This step includes the coding of the app based on the design. Android Studio, an open-source program for Android software development, will be used to develop the proposed mobile app. The iteration process has 3 phases, where a patient module and a physician module will be designed, and integration with an HIS will be performed as depicted in Figure 2.

**Step 4: Implementation**
After development, each module of EDDy will be pilot-tested for its intended use. The modules will be connected to a web-based backend database application to help the validation of content, visualization, and analysis of the data collected.

**Step 5: Testing and Quality Assurance**
The developed app will undergo a standard testing and quality assurance process to identify and rectify any issues and bugs. Each module will be validated during testing to ensure its intended performance.

**Step 6: Analysis of Feedback**
A pilot test among 2 potential users will be carried out at each level of iteration, and user experience will be measured using a Likert scale on various aspects of the app’s functionality. The parameters that will be measured during the piloting process will be the app’s functionality, usability, accessibility, compatibility, and performance under real-life scenarios.

Once the app is fully developed and interfaced with an HIS, it will be installed in the respective systems of the stakeholders and tested for its performance. Gaps and technical glitches will be rectified at each phase.

**Step 7: Deployment**
The final app will be deployed among end users (patients and physicians).

**Step 8: Maintenance**
The final app will be integrated into an HIS. Once the module is fully functional, the study will proceed to phase 3.

**Expected Outcomes From EDDy**
The app is expected to improve the SMBG practices in patients with T2DM. EDDy is expected to motivate the use of the app (acceptability) and improve the user experience on the contextual relevance. The ease of using the customized system and a clinician’s feedback mechanism through the physician
module is expected to improve the clinical utility of the app. It is expected that adherence to using the app would improve compliance to treatment protocol, thus improving treatment outcomes.

The entry of information into EDDy can be done either by the patient or their caregiver. The entry of SMBG values by the patient is a standard clinical practice in patient management. The current practice is that these self-monitored blood glucose values (SMBG data) is maintained by the patient on a printed form provided by the hospital. It is standard practice that such SMBG data function as a mechanism that helps physicians to optimize medication in routine clinical practice [10].

A researcher will provide instruction to the patient or patient party on how data entry needs to be done on the app. Additionally, a built-in user manual will be provided with the system. This will be developed as a part of the app development. A multilingual user manual or instructional videos will be developed to assist the patient in using EDDy. The pictogram in Figure 3 represents the workflow of EDDy.

Figure 3. Prototype of EDDy.

Phase 3
Study Design
Phase 3 will involve testing the adherence and usability of the developed app. This phase will use a pre- and posttest design among 48 patients with diabetes selected through purposive sampling.

At a 5% level of significance with 80% power and an effect size of 0.25, the minimum sample size required for a single-group, repeated measures study design involving 1 primary quantitative outcome (adherence) measured across 3 time points (preintervention, after 3 months, and after 6 months; correlation among repeated measures is 0.3) to understand the adherence to EDDy integrated with an HIS is 38 individuals. Incorporating a dropout rate of 20%, the minimum sample size requirement is 48 individuals.

The sample size was computed using G*Power (version 3.1.9; Universität Düsseldorf). The sample size may be amended (if required) based on pilot study findings.

The study setting will be the endocrinology OPD of a tertiary care hospital in coastal Karnataka, India.

Regarding study participants, patients with T2DM who have regular follow-up records will be considered for the study.

This study will analyze the adherence of patients with T2DM to EDDy in comparison to the existing manual method of documentation. Adherence will be evaluated in terms of the following:

- Adherence of entering SMBG values as predetermined by the patient: The study participants will be instructed to fix the frequency that they will monitor their blood glucose; based on this frequency, they will be prompted to conduct the glucometer testing and enter the values into EDDy. The values will be stored in the app and will be transferred to the server on a regular basis as the EDDy is integrated with an HIS.
- Adherence to diabetes self-management (DSM) practices: Pre- and posttest adherence to DSM practices will be analyzed using the same content-validated questionnaire.
- Adherence to routine blood investigations and screening regimen: Pre- and posttest adherence to periodic blood investigation and comorbidity screening will be analyzed.

The inclusion and exclusion criteria for sample collection will remain same as that of phase 1.
Phase 3 will be evaluated in 3 intervals:

- First evaluation (preintervention): During enrollment into the study, the patients will undergo an entry survey using a content-validated questionnaire to analyze the current DSM and SMBG practices they follow.
- Second evaluation (follow-up): An interim follow-up will be conducted in the third month.
- Third evaluation (postintervention): The same questionnaire used in the first evaluation will be used to evaluate DSM and SMBG practices after the intervention.

The study participants will be provided with EDDy on their mobile phones and will be trained to use it along with standard diabetes care. Training materials will be a part of the app for patient’s ready reference. The pre- and posttest process is shown in Figure 4.

![Figure 4. Pre and Posttest Process of development of EDDy.](image)

**Analysis**

Study participants will be followed up for 6 months and will be analyzed primarily for their adherence to using EDDy. It will be assessed based on patients’ willingness to use it, the ease of use, their acceptance of EDDy in terms of the contextual relevance, and the clinical utility. The participants will undergo an exit survey at the end of the study period using a validated questionnaire. It will be specific for assessing the qualitative parameters such as *user experience*, *glycemic control*, *adherence leading to improvement in their SMBG*, and *clinical utility*.

The qualitative outcomes are as follows:

- **Ease of use and adherence**: User’s perception on accessing the various modules in the app, ease of entering information, ease for retrieving information, and ease of comprehending information
- **Glycemic control**: Adherence in terms of regular monitoring and tracking of blood glucose values to keeps it under check, leading to well-managed T2DM
- **Contextual relevance**: Relevance of the contents of EDDy in terms of the user expectations and requirements as expressed in phase 1 of the study
- **Adherence to SMBG and DSM practices**: Reminders and alerts that ensure adherence to medications, insulin injections, regular glucose testing, and comorbidity screening
- **Clinical utility**: Improvement in treatment outcomes due to adherence to reminders or alerts and monitoring of blood glucose, adherence to medication, and follow-ups

The frequency will be tested based on the Diabetes Co-Conditions Screening Checklist provided by the Association of Diabetes Care & Education Specialists [11].

The following standard tools will be used for assessment, which are modified according to the study requirements and will be content validated by experts:

- **Perceived Health Web Site Usability Questionnaire**: A 12-item questionnaire that is validated for assessing the usability of health websites for older adults will be used [12].
- **Diabetes Self-Management Questionnaire**: It describes self-care activities related to diabetes. SMBG practices will be analyzed using this tool. A validated 12-item questionnaire to assess self-care activities associated with glycemic control will be used [13].
- **A self-developed, content-validated questionnaire**: Adherence will be evaluated, including the regularity of use; ease of use with respect to entering diabetes-related information such as regular testing and entry of glucose levels; adherence to medications or insulin doses; reporting of diabetes-related emergencies or complications, if any; and adherence to periodic screening such as laboratory
investigations, peripheral neuropathies, retinopathy, diabetic foot, and BMI.

The physicians will be interviewed to assess the clinical utility of the app and their experience in using the app using a content-validated questionnaire.

**Expected Outcomes**

A customized mHealth app that provides adequate information specific to the T2DM management will be developed. This can lead to an effective intervention that will reduce the risk of a person diagnosed with T2DM from developing further disease complications, such as chronic kidney disease, diabetic foot, retinopathies, and peripheral neuropathies, thereby reduce the cost of health care expenditure.

**Results**

**Pilot Study**

Out of 50 patients with T2DM (aged 40-50 years), 22 (44%) documented their self-monitored glucose values, whereas the remaining 28 (56%) did not. The percentage score of sources used by patients with diabetes to document the values of self-care activities showed that most of the patients (n=21, 42%) were using a personal diary to document, whereas only 1 (2%) patient each used a diabetes app and mobile notes. Responses from patients with diabetes regarding the necessity of personal health information in diabetes management showed that 31 (62%) patients were not completely aware of personal health information management.

The study results reflected the perception of patients with T2DM about their personal health information documentation practices. Additionally, the study gave a brief outlook about the challenges and barriers that patients face during their DSM activities.

**Study Status**

Phase 1 was completed on November 28, 2023. Phase 2 of the study, which is the app development, commenced in December 2023 and will end by March 2025. The wireframe model of EDDy is shown in Multimedia Appendix 1. Phase 3 of this study will follow afterward, analyzing the effectiveness of adherence to the app and its usability.

**Discussion**

**Pilot Study Findings**

It was evident from the pilot study that there is a strong requirement to promote documentation of SMBG and DSM practices. Overall, the study concluded that there is a scope for strengthening DSM, knowledge development, and improved self-documentation practices by promoting personal health information practices among patients with diabetes. It has the potential to improve the quality of life of patients with diabetes by better promoting self-care.

**Expected Findings**

When considering a country such as India, which has a very diverse population belonging to different socioeconomic backgrounds, it is difficult to implement a standard practice to tackle a problem. Statistics show that out of 45% of Indians over 45 years of age, about 11.5% are diagnosed with diabetes or high blood sugar levels [3]. Patients belonging to this age group are comparatively less familiar with technology and digital apps. Understanding their requirements and developing a hospital-based electronic diary may promote its use and adherence among them and, eventually, all patients with T2DM.

Considering India’s socioeconomic background, the requirements for this population need to be assessed from various angles. It can be demonstrated by analyzing the findings of the study conducted by Walle et al [14] among 422 patients with diabetes in Ethiopia. The study aimed at understanding the willingness of patients with diabetes mellitus to use mHealth app and its associated factors for self-care management in a low-income country as a precursor to digital health implementation in Ethiopia. The study concluded that the mHealth app developers should consider factors such as the patients’ age, place of residence, internet connectivity, attitude, perceived ease of use, and perceived usefulness while developing similar apps [14].

The pilot study conducted to analyze the awareness of patients with T2DM also clearly stated that they were aware of the importance of SMBG and DSM practices and the availability of mHealth apps to help them manage their T2DM. However, they were not very keen on using it due to various personal factors.

It was evident from the findings that there is a clear requirement to generate awareness about T2DM management and the use of mHealth apps. It is equally important to understand the challenges that keep them from using available mHealth apps. Based on these observations, if an mHealth app can be developed and the information can be transferred to the patients’ records in real time, it will be beneficial for both the patients and the doctors to better manage the condition.

**Conclusion**

The literature review and pilot study conducted suggests that an easily manageable, personalized diabetes-monitoring system for patients to do regular self-checks and that provides the end users with accurate data on the patient’s diabetic history will be beneficial for better management of T2DM. It would enable better sharing of information between patients and doctors, thus improving the communication between patients and doctors, which is significant in the management of diabetes and leads to better treatment outcome. The available apps are not very popular among the end users as they find it difficult to use and understand the contents. To ensure better usability and adherence, a convenient and simple alert system that enables the patient to conduct glucose testing at intervals set by the patients themselves can be designed. The proposed app EDDy will provide grading options to enter the diabetic-related complications (such as neuropathy, retinopathy, diabetic ulcers, etc), thus alerting patients and doctors for prompt action. It is also expected to enhance patients’ knowledge of tracking and managing the complications of diabetes and help in maintaining the visual representation of glucose values and complications. The simplicity and usability of the module are its novelty, which
may motivate the patients to keep track of their glucose values and help them to attain better health outcomes.

Acknowledgments
The author would like to acknowledge the participants of the pilot study and Mr Prajwal L Salins and Ms Ashika Shetty.

Data Availability
As this is a study protocol, the results of the pilot study have been reported, and the data will be submitted as required.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Wireframe model of Electronic Diabetes Diary (EDDy).

References
Abbreviations

**CTRI:** Clinical Trial Registry of India  
**DSM:** diabetes self-management  
**EDDY:** Electronic Diabetes Diary  
**IDI:** in-depth interview  
**IEC:** Institutional Ethics Committee  
**HIS:** hospital information system  
**mHealth:** mobile health  
**OPD:** outpatient department  
**SMBG:** self-management of blood glucose  
**T2DM:** type 2 diabetes mellitus

©Geena Skaria, Bhageerathy Reshmi, Sabu K M, Sahana Shetty, Vani Lakshmi R. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.01.2024. 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 Effects of the Processing of Positive Memories Technique on Posttrauma Affect and Cognitions Among Survivors of Trauma: Protocol for a Daily Diary Study

Talya Greene1,2*, PhD; Ateka A Contractor3*, PhD; Sheila Daniela Dicker-Oren2, MA; Andrea Fentem3, MA; Sharon R Sznitman4, PhD

1Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom
2Department of Community Mental Health, University of Haifa, Haifa, Israel
3Department of Psychology, University of North Texas, Denton, TX, United States
4School of Public Health, University of Haifa, Haifa, Israel
*these authors contributed equally

Corresponding Author:
Talya Greene, PhD
Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place
London, WC1E 7HB
United Kingdom
Phone: 44 020 7679 1897
Email: talya.greene@ucl.ac.uk

Abstract

Background: The Processing of Positive Memories Technique (PPMT) is a promising new treatment approach for posttraumatic stress disorder (PTSD), which involves detailed narration and processing of specific positive autobiographical memories. Indeed, preliminary case-series studies have found reductions in PTSD symptoms, negative affect, and negative cognitions among survivors of trauma who have received PPMT. However, PPMT’s effects have not been investigated at the daily level. In this study, we describe the protocol for a study that will examine the daily-level impacts of PPMT in a trauma-exposed, nonclinical community sample.

Objective: This study uses an innovative research protocol that combines case-series design and daily diary approaches to examine changes in daily affect, daily cognitions, and daily PTSD symptoms pre- and post-PPMT. We hypothesize that at the daily level, in comparison to their own pre-PPMT levels, following the PPMT intervention, participants will report (1) a lower count of endorsed daily PTSD symptoms, (2) increases in daily positive affect and decreases in daily negative affect, (3) increases in positive affect reactivity to daily positive events, and (4) decreases in daily posttrauma cognitions.

Methods: We are currently recruiting participants (target n=70) from a metroplex in the southwest United States. Following a screening survey, eligible participants complete a preintervention baseline survey, followed by 21 daily surveys in their natural environments. Then, they receive 4 PPMT sessions on a weekly basis. After the conclusion of the PPMT intervention, participants complete a postintervention outcome survey and 21 daily surveys. To compare daily affect, daily cognitions, and daily PTSD symptoms before and after PPMT, we will use the daily diary report data and conduct multilevel random intercepts and slopes linear regression models.

Results: Data collection was initiated in March 2022 and is expected to end by June 2024. As of November 28, 2023, a total of 515 participants had consented to the study in the screening phase. No analyses will be conducted until data collection has been completed.

Conclusions: Study findings could clarify whether deficits in positive autobiographical memory processes may also characterize PTSD alongside deficits in traumatic memory processes. Furthermore, PPMT could be an additional therapeutic tool for clinicians to help clients reduce posttraumatic distress in their everyday lives.

International Registered Report Identifier (IRRID): DERR1-10.2196/51838

https://www.researchprotocols.org/2024/1/e51838
Introduction

Posttraumatic stress disorder (PTSD) is a psychological condition that can develop following exposure to traumatic events, characterized by intrusive thoughts and re-experiencing symptoms, avoidance of trauma reminders, negative cognitions and mood, and heightened arousal [1]. A proposed mechanism underlying the development and maintenance of PTSD is disruptions in the encoding, consolidation, and retrieval of negatively- and positively-valenced autobiographical memories [2-4]. Evidence suggests that survivors of trauma with PTSD report difficulties accessing and detailing positive autobiographical memories [5-7], akin to reported difficulties with traumatic autobiographical memories [8]. Unsurprisingly, PTSD interventions typically target the content, processes, and phenomenological characteristics of autobiographical memories [9,10]. While such trauma-focused interventions (eg, prolonged exposure and cognitive processing therapy) are effective for many survivors of trauma, a significant proportion of survivors of trauma do not respond to these treatments, and there is a high degree of dropout from these interventions [11]. This highlights the need to develop alternative therapeutic approaches.

While most PTSD interventions address engagement with only traumatic autobiographical memories, only a few other interventions address both traumatic and positive autobiographical memories or only positive autobiographical memories and have been shown to be effective. For instance, Memory Specificity Training (targeting one’s ability to retrieve specific autobiographical memories irrespective of valence) is effective for PTSD and posttrauma distress [12,13]. Also, Broad-Minded Affective Coping, a positive emotion induction technique through the retrieval of positive autobiographical memories, improves mood among individuals with PTSD [14]. Such evidence suggests that a focus on positive autobiographical memories may be a helpful target in PTSD interventions.

The PTSD-Positive Memory Model [15,16] outlines that when survivors of trauma repeatedly retrieve, relive, and detail specific positive autobiographical memories, they may experience an improvement in PTSD symptoms, affect, and beliefs over time [15]. This model was foundational to the development of the Processing of Positive Memories Technique (PPMT), which is a 4- to 5-session intervention tailored to PTSD symptoms. During PPMT sessions, survivors of trauma are guided to narrate details of salient positive autobiographical memories; to access and strengthen positive values, affect, strengths, and thoughts associated with these memories; and to engage in positive affective, cognitive, and behavioral changes [17]. PPMT is influenced by positive psychology, a field that emphasizes factors and mechanisms that enhance psychological well-being rather than focusing solely on pathology [18,19]. PPMT draws from positive psychology interventions (eg, sharing positive narratives with others and using mental imagery to re-experience positive events) [19-21] and from interventions that increase memory retrieval to improve mental health [22]. The detailed session-by-session content of PPMT is outlined by Contractor and colleagues [17].

Practicing PPMT may help survivors of trauma retrieve more positive autobiographical memories over time, which may also translate to retrieving fewer negative autobiographical memories. Consequently, survivors of trauma may be able to better contextualize and integrate traumatic autobiographical memories with existing beliefs [23] and with other memories [2,24], which in turn could aid recovery after a trauma [2,25]. Positive autobiographical memories may also become primary reference points to interpret experiences and influence self-concept [26-28]. Furthermore, by repeatedly retrieving positive autobiographical memories and associated content, survivors of trauma may lessen their focus on negative material, experience more positive affect, and downregulate negative affect [29-32]. This may be especially helpful for survivors of trauma who experience emotional distress from retrieving negative autobiographical memories. In turn, this improved affect may help survivors of trauma positively interpret events [29,33] and note more positive content in their thoughts [34]. Overall, retrieving positive autobiographical memories may improve well-being [35], resilience [36], and adaptive coping [37], serving as a reminder that there are positive values and thoughts to hold on to despite the hardships faced by survivors of trauma.

Pilot studies have shown that PPMT is feasible and may improve therapeutic outcomes for survivors of trauma. Using an experimental design, a 2-session modified-PPMT [38] and a 5-session PPMT protocol [39] were compared to a neutral memory condition among survivors of trauma. In the first study, authors found that participants who repeatedly narrated the content of positive autobiographical memories reported decreases in PTSD symptom severity and negative affect, as well as increases in positive affect across time compared to the control condition [38]. In the second study, authors found that survivors of trauma who repeatedly retrieved positive (and neutral) memories reported less PTSD and depression severity, fewer posttrauma cognitions, and improved affect [39]. Using an open-label pilot trial, the feasibility and effects of the 5-session PPMT were examined among 12 survivors of trauma [40,41]. The authors found that PPMT reduced PTSD symptoms, reduced negative affect, and improved regulation of positive affect, and there were good feasibility indicators for PPMT (eg, PPMT was acceptable).

Critically, no study has examined PPMT’s effects using a larger community sample, nor has there been any exploration of whether PPMT is associated with postintervention changes in how survivors of trauma react to events in daily life. We can hypothesize that PPMT may impact individuals’ daily-life affect and cognitions; these impacts represent hypothesized
mechanisms through which PPMT may reduce PTSD symptom severity over time. Most studies examining PTSD intervention impacts use case-series designs, in which data are collected from a group of individuals pre- and postintervention and an aggregate assessment of symptomatology pre- and postintervention is conducted and compared. However, this approach has some noteworthy limitations. Affect, cognitions, and PTSD symptoms are dynamic and vary daily in response to trauma reminders and experiences [42-44]; thus, evaluating intervention effectiveness using 2 snapshot assessments of symptomatology is not sufficiently reliable or nuanced. Furthermore, case-series designs do not examine the daily-level mechanisms of change for an intervention. Given that PPMT may impact daily life affect, cognitions, and symptoms, it is crucial that the data enable an examination of these constructs at the daily level.

These limitations can be overcome by integrating case-series designs with a daily diary framework for the pre- and post-PPMT assessments. Daily diary studies are an intensive longitudinal data collection method in which participants provide daily reports of their experiences each day over a period of time. Compared to retrospective assessments, daily diary data are considered more ecologically valid as they are collected in an individual’s everyday life rather than in a laboratory, more accurate and robust, and less vulnerable to recall bias [45]. Thus, the proposed study outlines information on the protocol of an ongoing study that combines case-series design and daily diary approaches to provide novel insights into PPMT’s effects on daily-level cognitive and affective experiences. This approach can be conceptualized as a special subtype of case-series design that enables an examination of within-person changes at the daily level.

Specifically, the proposed study aims to use daily diary data pre- and post-PPMT to examine changes in daily PTSD, daily affect, and daily cognitions. We hypothesize that at the daily level, participants will report (1) a lower count of endorsed daily PTSD symptoms pre- to post-PPMT, (2) increases in daily positive affect and decreases in daily negative affect pre- to post-PPMT, (3) increases in positive affect reactivity to daily positive events (a within-person index of linear relations between daily positive events and daily positive affect) pre- to post-PPMT, and (4) decreases in daily posttrauma cognitions pre- to post-PPMT. The aim of this study is to detail the proposed research protocol and our hypotheses regarding daily PTSD, daily negative affect, and daily cognitions pre- and post-PPMT. As a supplementary analysis, we will examine if participants report a greater count of retrieved specific positive memories pre- to post-PPMT.

**Methods**

**Study Design**

The study involves four phases: (1) screening phase (eligibility survey), (2) preintervention phase (baseline survey and daily surveys), (3) intervention phase (PPMT and weekly surveys), and (4) postintervention phase (outcome survey and daily surveys). Figure 1 provides an illustration of the study procedure. All assessments are completed by participants using a computer or smartphone. Greene and colleagues [46] provide a detailed protocol for the questionnaires. Briefly summed up, following a web-based screening survey, eligible participants are asked to complete a preintervention phase baseline survey followed by 21 daily surveys in their natural environments. Then, they receive 4 PPMT sessions, completing 1 web-based survey per session and a feedback survey in the last session. After the conclusion of the PPMT intervention, participants are asked to complete a postintervention phase outcome survey and 21 daily surveys.

**Participants**

Participants (target n=70) are currently being recruited from a metroplex in the southwest United States through social media postings, flyers at businesses and public places, and university announcements since March 2022. The inclusion criteria are: (1) being aged between 18 and 65 years; (2) endorsing a trauma with posttrauma symptoms assessed by the Primary Care PTSD Screen for DSM-5 [47]; (3) access to an electronic device (eg, a computer or a smartphone) with internet capabilities; (4)
working knowledge of English; (5) no active suicidal plan, suicidal attempt, homicidal plan, or homicidal attempt (past 3 months including current); (6) being a current resident of the Dallas Fort Worth metropolex; (7) not currently in therapy with a mental health provider; (8) willingness and availability to participate in approximately 10 weeks of this study (including 4 therapy sessions); and (9) willingness to be video-recorded during sessions for quality control purposes.

Procedure

Screening Phase

During this phase, interested participants complete an eligibility survey. First, they read the informed consent document (information about the study, eligibility criteria, compensation, benefits and risks to study participation, and steps to ensure data confidentiality) and provide electronic informed consent if they wish to participate. Next, they answer questions to determine eligibility. Eligible and consenting participants are automatically redirected to a separate survey, wherein they provide contact information for study purposes. Research personnel then contact eligible participants to provide more information on the study (eg, survey timelines and PPMT sessions). No compensation is provided for this study phase.

Preintervention Phase

Participants complete 1 baseline survey and 21 daily surveys as part of this phase. The baseline survey contains questions on demographics, trauma history, PTSD symptoms, and other psychological symptoms, as well as affect and cognitive processes (approximately 30-minute completion time). Eligible participants receive the baseline survey link by email at a date determined to be feasible based on contact with participants. Participants are given up to 48 hours to complete the baseline survey and are sent reminders if they do not complete the survey in a timely manner. Research personnel monitor survey responses for completion, response times to identify unfeasibly short times, accurate participant ID entry, therapy history to confirm they are not currently in therapy, and trauma history.

Participants who complete the baseline survey are asked to complete 21 daily surveys. The daily surveys include questions assessing daily PTSD symptoms, affect, cognitions, and events that occurred in the last 24 hours (approximately 3-5 minutes to complete each survey). The link to the first daily survey is emailed to participants within 1-3 days after completing the baseline survey. Participants receive daily surveys once a day at fixed intervals (at 7:00 PM each day) over a 21-day period, and they have until 11:59 PM to complete each daily survey. Participants are sent text reminders for survey completion to enhance compliance and are contacted if they miss any surveys.

Intervention Phase

PPMT is administered weekly as a 4-session protocol during the intervention phase. The sessions are scheduled within 1-2 weeks after completing the preintervention phase. In session 1, participants receive psychoeducation on PTSD symptoms, an overview of PPMT, and are assessed for psychological symptoms. Sessions 1-4 involve the detailed processing of a salient positive autobiographical memory to elicit “values, affect, strengths, and thoughts” related to that positive memory. Homework assignments include listening to an audio recording of that memory, completing a “values, affect, strengths, and thoughts” log, and engaging in a behavioral activity. In session 4, the therapist also reviews psychological symptoms and addresses termination. Following the completion of session 4, the participants complete a feedback survey on PPMT.

Postintervention Phase

Participants complete 1 outcome survey and 21 daily surveys as part of this phase. Within approximately 1 week after completing the intervention phase, participants complete an outcome survey. The procedural aspects of this phase mimic the preintervention phase. The outcome survey has questions similar to those of the preintervention phase baseline survey (without demographics and trauma history items and with different cue words for the measure examining the count of retrieved memories).

Dropout

Participants who do not complete the outcome survey after 2 days, consecutively miss 4 daily surveys, or miss more than 8 daily surveys (<60% of the daily surveys) in either the preintervention or postintervention phases are considered dropouts for this study, and they do not continue to receive survey links in order to avoid burdening participants with repeated requests to complete the surveys if they no longer wish to participate.

Study’s Primary Measures

Overview of Primary Measures

In this section, we outline the measures that relate to the primary outcomes of this study. The primary outcomes of interest are daily positive affect levels, daily positive affect reactivity (within-person index of linear relations between daily positive events and daily positive affect [48]), daily negative affect levels, daily posttrauma cognitions, daily PTSD symptoms, and the number of retrieved specific positive memories as measured pre- and postintervention. Secondary measures (eg, difficulties in positive emotional regulation using the “Difficulties in Emotion Regulation Scale-Positive” and the severity of PTSD symptoms using the “PTSD Checklist for DSM-5”) are also administered to allow for the assessment of possible mediators and moderators of treatment effects. Table S1 in Multimedia Appendix 1 [40,47,49-58] provides detailed information on all study measures (including measures for supplemental analyses).

Preintervention Phase Baseline Survey

The number of retrieved specific-positive memories is measured by the Autobiographical Memory Test (AMT) [49,59]. The AMT uses a cued memory recall technique involving the presentation of individual cue words, followed by a prompt to recall a personally meaningful and specific memory of an event that took place within any 24-hour period. For this study, participants are shown cue words and asked to retrieve a personal and specific memory of the cue-word-related event within 60 seconds [59]. The instructions were adapted from previous autobiographical memory studies [49,60,61]. In the preintervention baseline survey, we included 5 cues drawn from...
previous studies: friendly, happy, honest, kind, and humorous [62-64]. We will follow coding guidelines to categorize AMT responses [65,66]. AMT responses will be coded as specific (event that occurred at a certain place within 24 hours), extended (event that lasted >1 day), or categorical (summary of repeated events). AMT responses will also be coded as positive or nonpositive following the Coding and Assessment System for Narratives of Trauma [67]. Lastly, AMT responses will be coded as semantic associate (no personal memory) and omission (did not retrieve the memory within 60 seconds or was unable to recall a memory). The AMT demonstrates good psychometrics [68].

Preintervention Phase Daily Surveys

Daily negative and positive events are measured by asking participants to rate their most positive and negative events in the last 24 hours from 0 (not at all unpleasant) to 3 (very unpleasant) [50].

Daily affect levels (ie, positive and negative) are assessed by rating the extent of 4 positive (excited, cheerful, satisfied, and relaxed) and 6 negative (stressed, irritated, anxious, sad, hopeless, and insecure) emotions in the last 24 hours. These emotions were used in a previous daily diary study based on a theoretical circumflex of emotions. The study showed excellent between-person reliability and good within-person reliability [69]. In this study, responses are rated on a 5-point Likert scale ranging from 1 (very slightly or not at all) to 5 (extremely).

Daily posttrauma cognitions are measured by the Brief Version of the Postransformative Cognitions Inventory [51], which is a 9-item self-report measure assessing postransformative cognitions. Responses are provided on a 7-point Likert scale ranging from 1 (totally disagree) to 7 (totally agree), with the time frame modified to “in the last 24 hours.”

Daily PTSD symptom severity is assessed using the Primary Care PTSD Screen for DSM-5 [47], which is a 5-item self-report measure with dichotomous “yes” or “no” items; the time frame is modified to “in the last 24 hours.” We will use the count of endorsed PTSD symptoms as a measure of daily PTSD symptom severity.

Postintervention Phase Outcome Survey

Similar to the preintervention phase baseline survey, the AMT is administered to examine the number of retrieved specific positive memories. Different cue words are used: peaceful, loyal, helpful, safe, and love [62-64].

Postintervention Phase Daily Surveys

The procedural aspects and measures mimic the preintervention phase daily surveys.

Participant Safety and Psychological Distress

There are minimal foreseeable risks associated with this study. Few participants may experience an increase in PTSD severity [70] or suicidal ideation [71], attributed to the sensitive nature of questions targeting trauma reactions and to PPMT itself. To address any such concerns, the study co-principal investigator (co-PI; a licensed clinical psychologist) is training all research personnel on considerations in administering assessments to survivors of trauma, providing clinical training and weekly supervision to study therapists, and training study therapists in anxiety-reduction techniques (eg, guided imagery). Furthermore, study participants receive information on mental health services.

To address any risk factors for self-harm, we are closely monitoring participants during PPMT to assess for any reported suicidal ideation, plan, or attempt (eg, administration of a depression measure every session). Furthermore, the study co-PI is training study therapists to conduct in-depth suicide risk assessments. In the event of suicidal ideation, plan, or attempt reported by participants in session, study therapists conduct a suicide risk assessment and consult with the supervising co-PI to determine if emergency services need to be contacted to ensure participant safety. Also, the participant or therapist may discontinue the study at any time should symptoms worsen or if the participant simply desires to withdraw. Lastly, we are providing information on community mental health centers offering 24-hour access to services and emphasizing contacting 911 or 988 in times of imminent risk.

Treatment Nonresponse or Relapse

Any treatment for postransformative mental health is associated with some chance of failure to respond or relapse [72]. We are implementing the following procedures to address any such potential concerns. If a participant shows substantial increases in PTSD or depression severity or reports risk factors for self-harm during intervention sessions, we are providing mental health referrals immediately for alternate treatment options, including a 24-hour access local mental health center. Furthermore, research personnel are contacting participants when they miss appointments to check on their health status.

Intervention Training

Therapist Training

The study co-PI trained doctoral students (ie, study therapists) in PPMT. This training included a review of PPMT’s theoretical underpinnings, manuals, and fidelity checklists, as well as practice in PPMT administration. Furthermore, study therapists are required to follow detailed session protocols during each PPMT session. Research assistants were trained in PPMT fidelity ratings; ≥0.81 kappa coefficient and ≥0.8% percent agreement will be considered acceptable interrater reliability (IRR) [73].

Treatment Delivery

All sessions are being video recorded. The co-PI has reviewed all recorded PPMT sessions for 1 participant for each study therapist; she will continue to review 20%-50% of the video-recorded sessions as needed [74]. The co-PI is providing weekly group supervision to study therapists that involves case discussions and feedback.

Fidelity Ratings Across Raters

The authors have created fidelity checklists that include a list of proscribed PPMT components to be recorded as occurring or not occurring. Using these fidelity checklists, 2 trained evaluators will independently code video-recorded sessions for 18-20 participants. These data will be used to compute IRR estimates. If acceptable IRR estimates are not achieved, trained
evaluators will code an additional 20% of sessions. Once acceptable IRR estimates are achieved, the evaluators will solely code the remainder of the treatment sessions for fidelity.

**Adherence to PPMT Components**

We will compute percentage adherence across sessions for each of the trained study therapists, with the recommended 80%-100% benchmark indicating high fidelity [74].

**Data Analysis**

**Power Analysis**

We conducted an a priori power analysis using the **EMAtools R** (Kleiman) package [75] for power curves for multilevel studies. The power analysis was based on two 3-week assessment bursts with 1 questionnaire per day and an estimated intraclass correlation coefficient of 0.36 based on a previous study on daily-level emotions, cognitions, and PTSD [76]. Analysis showed that 70 participants and up to 25% missing data would be sufficient to detect a medium effect size \(d=0.5\) with 80% power.

**Analytical Plan**

A paired sample \(t\) test will be used to examine changes in the count of retrieved specific-positive autobiographical memories pre- versus post-PPMT (comparing the preintervention baseline and postintervention outcome surveys). To examine changes in daily affect, daily cognitions, and daily PTSD symptoms, we will use the daily diary reports pre- and postintervention and conduct multilevel random intercepts and slopes linear regression models for each outcome variable using **MPlus 8.3** (Muthén and Muthén), **nlme** (Pinheiro et al), and **lme R** (Bates et al) packages, comparing the models pre- and post-PPMT with and without demographic covariates (eg, gender, age, and education). We will also conduct exploratory analyses examining additional variables included in the study as predictors or moderators of post-PPMT outcomes (eg, count of trauma types previously experienced, PTSD severity at baseline, and difficulties in positive emotional regulation).

**Missing Data**

At the survey level, the web-based questionnaire has been set up with a prompt if questions have been skipped, with the option to continue the survey without completing a particular item or to go back and complete the skipped question. We anticipate a little missing data within the submitted surveys. We will treat surveys that have been submitted with missing data as complete for the purposes of determining dropout and participation compensation. An analysis will be conducted to investigate the pattern of missingness. If data are missing at random or missing completely at random, they will be handled by listwise deletion and models fit by maximum likelihood. If data are not missing at random, then missing data will be imputed using the **MICE** (van Buuren et al) package in R [77] and fit by maximum likelihood.

**Ethical Considerations**

The institutional review boards at the University of North Texas (#21-420) and the University of Haifa (#480/21) have approved this study. During the screening phase, interested participants read the informed consent document (information about the study, eligibility criteria, compensation, benefits and risks to study participation, and steps to ensure data confidentiality) and provide electronic informed consent if they wish to participate. Participants are then contacted by research personnel to re-explain any study procedures and obtain or confirm identifying information.

In terms of compensation, participants receive US $1.50 for each completed daily survey and US $10 each for completing each of the baseline and outcome surveys. Participants receive US $10 for completing each of the 4 PPMT session surveys and US $12 toward transportation costs cumulatively for all 4 (attended) intervention sessions. In order to incentivize participants to provide as much data as possible, participants who complete 36 surveys without any missing data receive an additional US $15. The total potential compensation for participation is US $150.

Participants provide personally identifiable information (eg, name and contact information), which is only used for scheduling purposes, study-related communications, and to connect data longitudinally. Each participant receives a unique and randomly generated ID number, which is used on all web-based surveys for this study. At no point is any personally identifiable information linked to participant data. Furthermore, deidentified data will be analyzed for the scientific dissemination of study findings.

**Results**

Year 1 of the study (October 2021—September 2022) was primarily devoted to recruiting and training research personnel, obtaining ethics approvals, and preparing to launch the study. Year 2 of the study (October 2022—September 2023) has been focused on participant recruitment and data collection. During Year 3 of the study (October 2023—September 2024), we will complete data collection from our targeted sample and start the data cleaning and analysis process. During Year 4 of the study (October 2024—October 2025), we will complete the data analyses and prepare planned scientific outputs (eg, publications and presentations).

Data collection was initiated in March 2022. As of November 28, 2023, a total of 515 participants had consented to the study in the screening phase. Of those, 258 (50.1%) participants were eligible, 92 (35.7%) of which gave their consent by phone and attempted the first daily survey of the study. Of the 92 participants who completed the first daily survey, a total of 58 (63%) participants completed all intervention sessions and ≥13 daily surveys (≥60% of the daily surveys), and 28 (30.4%) participants dropped out in various stages of the study, mostly before the PPMT intervention. As of November 29, 2023, a total of 6 (6.5%) participants are currently participating in various phases of the study. Data collection is expected to end by June 2024. No analyses will be conducted until data collection has been completed.
Discussion

This study aims to examine the daily-level impacts of PPMT, a promising adjunct or alternative to traditional PTSD treatments, in a trauma-exposed, nonclinical community sample. This study combines a case-series design and a daily diary design to examine potential mechanisms of change in PTSD symptoms by assessing daily affect, daily cognitions, and daily PTSD symptoms before and after the PPMT intervention. This approach enables a more nuanced and ecologically valid exploration of changes as compared with retrospective aggregate assessments. We outline the research protocol for this study, including the hypotheses and the proposed analyses. When data collection has been completed (estimated date: June 2024), we will test our hypotheses that, at the daily level, in comparison to their own pre-PPMT levels, following the PPMT intervention, participants will report (1) a lower count of endorsed daily PTSD symptoms, (2) increases in daily positive affect and decreases in daily negative affect, (3) increases in positive affect reactivity to daily positive events, and (4) decreases in daily posttrauma cognitions.

The findings of this proposed study could have significant implications. Results could clarify whether deficits in positive autobiographical memory processes (eg, retrieval and encoding) may also characterize PTSD alongside deficits in traumatic memory processes [7]. If the study’s hypotheses are confirmed, PPMT could be an additional therapeutic tool for clinicians to help clients with posttraumatic distress. Unlike other trauma interventions, PPMT exclusively targets positive autobiographical memories in treatment while redirecting attention away from negative content embedded in the positive memories. By uniquely combining positive and symptom-focused techniques and theories, PPMT aims to increase positive elements (eg, values, affect, and thoughts) while simultaneously decreasing PTSD severity [17].

There are some limitations to this study that should be considered. The study uses a self-report approach, which, although it reduces recall bias, is still subject to potential difficulties in the recall of experiences over the course of each day. Relatedly, while it reduces participant burden, the PC-PTSD measure is a PTSD symptom screener and usually is coupled with a comprehensive structured diagnostic interview for PTSD or a self-report measure assessing all 20 PTSD symptoms (which we do not do in this study due to the daily-level methodology and associated participant and time burdens). Furthermore, we are not gathering information on trauma characteristics such as the ages at which the trauma was experienced, the frequency or chronicity of each experienced trauma, or the time since the trauma has elapsed. Such information can impact posttrauma distress [78-80] and may moderate the impacts of PPMT in this study; hence, it should be empirically investigated in future research. In addition, while the research design of one assessment per day enables examination of changes in daily symptoms, affect, and cognition before and after the intervention, it does not have a sampling frequency nor sufficient power to examine fine-grained dynamic interactions between symptoms, affect, and cognition using even more complex modeling techniques. Lastly, our eligibility criterion permits individuals endorsing even 1 PTSD symptom at a clinical level to be included in the study. While such an approach accounts for impairment among individuals endorsing sub-threshold PTSD [81], it may also make it statistically difficult to detect any changes in PTSD symptoms (ie, the floor effect).

In conclusion, this study will contribute to the development of more personalized and alternative PTSD interventions for survivors of trauma who drop out or do not benefit from existing PTSD treatments. Further studies could examine PPMT as an ecological momentary intervention, wherein individuals receive daily reminders and instructions for engaging in therapy-relevant behaviors (eg, processing of positive memories). Such interventions can particularly benefit communities that do not have easy access to mental health services and are underserved in that regard [82]. Finally, this study will give insight into the mechanisms of change in the PPMT intervention through elucidating daily-level changes in affect, cognition, symptoms, and event reactivity.

Acknowledgments

This research was supported by a grant (2020017) from the United States-Israel Binational Science Foundation (BSF). The authors would like to acknowledge the work of Gurleen Kaur for participant recruitment and data collection and the work of Sidonia Compton as a study therapist.

Data Availability

The data sets that will be generated through this study will not be publicly available due to the sensitive nature of the posttrauma daily-level data, however the corresponding author will make deidentified data available on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplemental table 1. Timeline information on measures.

[DOCX File, 25 KB - resprot_v13i1e51838_app1.docx ]
References


46. Greene T, Contractor AA, Oren SDD, Fentem A, Szmitan S. Daily-level impacts of the processing of positive memories technique on post-trauma affect and cognitions among trauma survivors: a daily diary study protocol. Open Science Framework. URL: https://osf.io/5a4z4/ [accessed 2023-12-21]


Abbreviations
- AMT: Autobiographical Memory Test
- co-PI: coprincipal investigator
- IRR: interrater reliability
- PPMT: Processing of Positive Memories Technique
- PTSD: posttraumatic stress disorder

©Talya Greene, Ateka A Contractor, Sheila Daniela Dicker-Oren, Andrea Fentem, Sharon R Sznitman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.01.2024. 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

Microfragmented Fat and Biphasic Calcium Phosphates for Alveolar Cleft Repair: Protocol for a Prospective, Nonblinded, First-in-Human Clinical Study

Diandra Sabrina Natsir Kalla¹,², MD; Salem Alkaabi¹,³, DDS; Abul Fauzi⁴, DDS; Andi Tajrin⁴, DDS; Rifaat Nurrahma¹,⁵, DDS; Werner E G Müller⁶, Prof Dr; Heinz C Schröder⁶, Prof Dr; Xiaohong Wang⁶, Prof Dr; Tymour Forouzanfar¹,⁷, Prof Dr; Marco N Helder¹, PhD; Muhammad Ruslin⁴, Prof Dr

¹Department of Oral and Maxillofacial Surgery/Oral Pathology, Amsterdam University Medical Centers and Academic Centre for Dentistry Amsterdam, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, Netherlands
²Department of Biochemistry, Faculty of Medicine, Hasanuddin University, Makassar, Indonesia
³Department of Oral and Maxillofacial Surgery, Fujairah Hospital, Ministry of Health, Fujairah, United Arab Emirates
⁴Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Hasanuddin University, Makassar, Indonesia
⁵Department of Prosthodontics, Faculty of Dentistry, Hasanuddin University, Makassar, Indonesia
⁶ERC Advanced Investigator Grant Research Group at the Institute for Physiological Chemistry, University Medical Center of the Johannes Gutenberg University, Mainz, Germany
⁷Department of Oral and Maxillofacial Surgery, Leiden University Medical Centre, Leiden, Netherlands

Corresponding Author:
Diandra Sabrina Natsir Kalla, MD
Department of Oral and Maxillofacial Surgery/Oral Pathology
Amsterdam University Medical Centers and Academic Centre for Dentistry Amsterdam
Vrije Universiteit Amsterdam, Amsterdam Movement Sciences
De Boelelaan 1117
PO box 7057
Amsterdam, 1007 MB
Netherlands
Phone: 31 20 444 1150
Email: sabrina.diandra@gmail.com

Abstract

Background: Biphasic calcium phosphates (BCP) may serve as off-the-shelf alternatives for iliac crest-derived autologous bone in alveolar cleft reconstructions. To add osteoinductivity to the osteoconductive BCPs to achieve similar regenerative capacity as autologous bone, a locally harvested buccal fat pad will be mechanically fractionated to generate microfragmented fat (MFAT), which has been shown to have high regenerative capacity due to high pericyte and mesenchymal stem cell content and a preserved perivascular niche.

Objective: Our primary objectives will be to assess the feasibility and safety of the BCP-MFAT combination. The secondary objective will be efficacy, which will be evaluated using radiographic imaging and histological and histomorphometric evaluation of biopsies taken 6 months postoperatively, concomitant with dental implant placement.

Methods: Eight patients with alveolar cleft (≥15 years) will be included in this prospective, nonblinded, first-in-human clinical study. MFAT will be prepared intraoperatively from the patient's own buccal fat pad. Regular blood tests and physical examinations will be conducted, and any adverse events (AEs) or serious EAs (SAEs) will be meticulously recorded. Radiographic imaging will be performed prior to surgery and at regular intervals after reconstruction of the alveolar cleft with the BCP-MFAT combination. Biopsies obtained after 6 months with a trephine drill used to prepare the implantation site will be assessed with histological and histomorphometric analyses after methylmethacrylate embedding and sectioning.

Results: The primary outcome parameter will be safety after 6 months’ follow-up, as monitored closely using possible occurrences of SAEs based on radiographic imaging, blood tests, and physical examinations. For efficacy, radiographic imaging will be used for clinical grading of the bone construct using the Bergland scale. In addition, bone parameters such as bone volume, osteoid volume, graft volume, and number of osteoclasts will be histomorphometrically quantified. Recruitment started in November 2019, and the trial is currently in the follow-up stage. This protocol’s current version is 1.0, dated September 15, 2019.

https://www.researchprotocols.org/2024/1/e42371
Conclusions: In this first-in-human study, not only safety but also the histologically and radiographically assessed regenerative potential of the BCP-MFAT combination will be evaluated in an alveolar cleft model. When an SAE occurs, it will be concluded that the BCP-MFAT combination is not yet safe in the current setting. Regarding AEs, if they do not occur at a higher frequency than that in patients treated with standard care (autologous bone) or can be resolved by noninvasive conventional methods (eg, with analgesics or antibiotics), the BCP-MFAT combination will be considered safe. In all other cases, the BCP-MFAT combination will not yet be considered safe.

Trial Registration: Indonesia Clinical Trial Registry INA-EW74C1N; https://tinyurl.com/28tnrr64

International Registered Report Identifier (IRRID): DERR1-10.2196/42371

(KEYWORDS)

microfragmented fat; calcium phosphate; bone regeneration; regenerative medicine; alveolar; bone grafting; bone; graft; alveolar cleft; surgery; surgical; perioperative; mouth; oral surgery; maxillofacial; jaw; oral pathology; oral; dentistry; dental; tooth; teeth; osteo; osteoconductive biphasic calcium phosphate; autograft; operation

Introduction

Alveolar cleft is defined as a bone gap in the primary palate from the nasal sill to the incisive foramen [1]. The defect occurs as a result of disruption of primary palate development between 4 and 12 weeks of gestational age, specifically in the frontonasal prominence [2]. The treatment protocol varies on the basis of the following factors: timing, surgical procedure, and grafting material. Secondary alveolar bone grafting (SABG) is the most preferred and successful method that is usually performed during the mixed dentition period (6–11 years), which allows the provision of support to teeth eruption and facial growth [1]. The iliac crest as a bone graft donor for alveolar cleft reconstruction has gained popularity since it was first introduced by Schmid [3] in 1954, and, in particular, for SABG procedures because it allows harvesting of large amounts of bone for alveolar cleft surgery [4]. Other bone graft sources include the cranium, tibia, and the mandibular symphysis [5]. However, several studies have reported risks of general postoperative complications using autografts, such as pain, prolonged hospital stay, and donor site–specific complications such as scarring, cutaneous nerve injury near the iliac crest, and hematoma after harvesting the cranial bone [6-9]. Therefore, alternative materials are being evaluated for alveolar cleft surgery.

Biphasic calcium phosphate (BCP) is a bioceramic that consists of 2 materials, hydroxyapatite (HA) and β-tricalcium phosphate, mixed in different ratios [10]. It is a biocompatible, easy-to-handle, safe material with a mineral composition comparable to that of human bone tissue [10]. BCP has been mixed in vivo and in vitro with autografts, inducing factors or cells to improve its osteoinductivity [11,12], also in the fields of dentistry and maxillofacial surgery [13-15]. Although calcium phosphate ceramic is not yet considered standard-of-care, it has been used for alveolar cleft reconstruction with satisfactory results [16], reportedly providing support for teeth eruption [17].

Adipose tissue is a source of mesenchymal stem cells, and adipose stem cells (ASCs) can be collected with minimum risk and discomfort from the buccal fat pad (BFP) [18]. The BFP surrounds the buccinator muscle and other superficial muscles such as the masseter, the zygomaticus major, and the zygomaticus minor [19]. Moreover, multiple studies have shown that the cell yield of ASCs per volume is at least 100-500 times that of mesenchymal stem cells in bone marrow aspirates [18,20]. Commonly, ASCs are prepared using enzymatic (collagenase) digestion which, however, is considered “more than minimal manipulation” of the cells by the US Food and Drug Administration and the European Medicines Agency [21]. An alternative method, which also takes considerably less time, involves processing the adipose tissue mechanically into microfragmented fat (MFAT) [22]. MFAT is reported to have similar or even higher secretory activity of regenerative growth factors and cytokines and pericyte content than an enzymatically derived stromal vascular fraction (SVF) [23]. In addition, the MFAT procedure can be applied even in regular hospitals because its harvesting and processing does not require a major invasive surgery, specialized equipment or expensive disposables, or good manufacturing practices–qualified cell culture expansion. Autologous application of MFAT has, among others, been used with success for clinical reconstructions in the maxillofacial area [24].

We hereby describe the protocol of a first-in-human clinical safety trial using BCP mixed with MFAT for alveolar cleft reconstruction. Our hypothesis is that the combination will be a safe, efficient, and effective alternative to conventional autograft since the osteoconductive BCP is supplemented by the regenerative capacity from the MFAT.

Methods

Study Design

This first-in-human surgical study can be classified as a “stage 1” study in accordance with the IDEAL (innovation, development, exploration, assessment, and long-term study) framework [25]. It is a single-center prospective clinical trial comprising 8 patients, assessing the safety of a combination of MFAT and BCP (BoneCeramic, Straumann) as bone graft material for alveolar cleft reconstruction. The BCP is a synthetic bone graft containing 60% HA and 40% β-tricalcium phosphate, a porosity of 90%, and an interconnected pore size of 100-500 μm. The BCP will be combined in a 1:1 cm³ ratio with MFAT prepared from the patients’ own BFP, which is processed with a 1.2-mm single-use sizing transfer Tulip Gen II Nanofat Kit.
(Tulip Medical). The primary end point will be set at 6 months. At each follow-up visit, adverse events (AEs) or serious AEs (SAEs) will be documented, and clinical assessments will be performed at time points specified in the Interventions section. After these 6 months, a bone biopsy sample will be taken using a hollow drill during dental implant preparation and subsequently processed for histological or histomorphometric analysis (Figure 1). Finally, a report on safety and proof of concept with regard to bone formation will be made and published.

Figure 1. Simplified diagram of the study protocol (adapted from the SPIRIT [Standard Protocol Items: Recommendations for Interventional Trials] checklist; Multimedia Appendix 1). BCP: biphasic calcium phosphate; MFAT: microfragmented fat.

Ethical Considerations
The clinical trial protocol was approved by the ethical committee of Hasanuddin University, Makassar, Indonesia (1063/UN4.6.4.5.31/PP36/2019) and registered in the Indonesian trial registry (INA-EW74C1N). Informed consent will be obtained from candidates or their parents or legal guardians who are willing to join the trial after being fully educated about the trial procedure. This study complies with the principles of the Declaration of Helsinki. In the informed consent form signed by the patients or their legal representatives, consent to publish their data in a deidentified manner is included.

Inclusion and Exclusion Criteria
Patients will be included on the basis of the following criteria [26]: being healthy male or female participants aged ≥15 years, having unilateral alveolar cleft without any previous history of grafting procedures, being categorized as normal healthy patients for anesthetic risk per the American Society for Anesthesiologists’ criteria, and having a normal blood count.

Patients will be excluded on the basis of the following criteria [26]: having poor oral hygiene with mouth plaques; having systemic disease; having systemic or local infection; having received chemotherapy, radiotherapy, immunosuppressives, or anticoagulants that may interfere with the healing process; having received bone growth–inducing factors, malnutrition, or active influenza; and being pregnant.

Interventions
Under general anesthesia and infiltration with lidocaine (1%) with epinephrine (1:100,000 dilution), the surgeon will identify the Stensen’s duct with a lacrimal probe and make an incision 2-3 cm below the duct [27]. A dissection penetrating the muscles and the superficial fascia will allow spontaneous herniation of the BFP [27]. This procedure will be carried out bilaterally on both cheeks in order to obtain approximately 3 cm³ of fat. After vasoconstrictor infiltration with epinephrine (1:100,000), a full mucoperiosteal flap spanning the first molar to the central incisor is lifted. After exposure of the full alveolar cleft and to separate the nasal layer from the oral mucosa, the tissue was meticulously dissected. Following the reflection of a palatal mucoperiosteal flap from either side of the cleft, the palatal tissues are elevated. The oronasal fistula is repaired cranially by elevating and suturing the nasal mucosa [4], thereby creating a pocket for BCP-MFAT deposition.

In parallel with the defect surgery, the harvested fat will be chopped into small pieces with a scissor and soaked in normal saline for 10-15 minutes. The normal saline then will be drained and the chopped fat will be processed into MFAT using 2 syringes (size 10 cm³) connected with the 1.2-mm single-use sizing transfer Tulip Gen II Nanofat Kit in accordance with the manufacturer’s protocol. MFAT will be mixed with BCP (Straumann Bone Ceramic) in a ratio of 1 g:1 cm³ until it reaches homogenous consistency. The BCP-MFAT mixture will be placed as a graft material into the alveolar cleft defect. If the defect is large and requires more bone graft, another mixture will be prepared with the same mixing ratio. If necessary, a membrane will be used to cover the grafted defect. Finally, the defect will be closed by suturing the palatal mucoperiosteal flaps using absorbable sutures with 3-0 vicryl sutures for mucosa and 4-0 vicryl sutures for nasal reconstruction. All patients will be prescribed antibiotics and analgesics postoperatively.

AE Assessment
Any change in the health of subjects will be documented in their medical history, and required medical care will be provided. Any unexpected physical or laboratory change, symptom, or disease that occurs in a treated patient who has been administered the graft will be documented as an AE. An AE will be graded in accordance with the World Health Organization’s classification [28] as either serious or nonserious based on its intensity. The Clavien-Dindo Classification of Surgical Complications will also be used in case of any incidence [29]. In the case of an SAE, a report will be made to the sponsor within 24 hours and to the ethical committee within
3 days from the date of onset. If the SAE concerns severe toxicity or infection associated with graft products, the trial will be terminated immediately.

Sample Size
This is a first-in-human phase I clinical trial aimed to obtain insight on the safety and feasibility of the treatment with the BCP-MFAT combination. We assume that no SAEs or AEs will occur, based on our clinical experience with other applications of MFAT and the well-proven safety of BCP. Upon consultation with a statistician, a sample size of 8 individuals is expected to be sufficient for this trial.

Recruitment
Patients will be recruited from an existing database of the Hasanuddin University Dental Hospital, from general practices of Hasanuddin Dental Hospital and in the area around Makassar. Thereafter, we will determine whether the candidates fulfill the study’s inclusion and exclusion criteria. Thorough assessment and training regarding the safety measurements at the research site at Hasanuddin University Dental Hospital will be performed prior to the trial by the ethical and surgical teams.

Since we did not want to enroll children in a safety study with this novel concept in clinical practice, we chose to only include older adolescent and adult patients, being themselves capable of decision-making. Within Indonesia, this age group is more common due to cultural and religious backgrounds causing abstinance from cleft surgeries.

The trial will be conducted at Hasanuddin Dental Hospital. All participants will be asked to sign an informed consent form after risks and possible complications of the procedure (eg, bleeding, infection, cheek asymmetry, parotid duct injury, possibility of facial nerve branches injury, and—although not likely—nonclosure) were appropriately communicated with the patient. Data will be handled and stored in a coded—that is, deidentified—format, so that data cannot be traced back to the patient without a decoding key, which is stored in a locked place and only accessible to the study’s principal investigator. Implants will be offered free of charge.

Randomization and Blinding
Since this trial comprises only 1 type of treatment, no randomization or blinding to the treatment is possible.

Data Collection and Access
The research team will be informed about the rules and their responsibilities. All members of the research team, which will collect the data in accordance with the evaluation table (Table 1), will receive training on how data collection should be performed. The data manager will document the data in a patient-coded manner (ie, each patient will be assigned a study-specific code under which the data will be stored in order to conceal their identity), which will subsequently be handed over to the clinical evaluators and investigators. The primary end point is set at 6 months.

<table>
<thead>
<tr>
<th>Table 1. Patients’ assessments.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Consent form</td>
</tr>
<tr>
<td>Preoperative</td>
</tr>
<tr>
<td>Operative day</td>
</tr>
<tr>
<td>Postoperative day 1</td>
</tr>
<tr>
<td>Postoperative day 8</td>
</tr>
<tr>
<td>Postoperative day 14</td>
</tr>
<tr>
<td>Postoperative day 30</td>
</tr>
<tr>
<td>Postoperative day 90</td>
</tr>
<tr>
<td>Postoperative day 180</td>
</tr>
</tbody>
</table>

Posttrial Care
After the primary end point assessment, the participants will be followed up for an additional period of 3 years to ensure their safety and to record whether any delayed side effect occurs as a result of treatment with the BCP-MFAT combination, as previously done in a similar study [26].

Monitoring
Internal monitors of the Ethics and Research Committee, Faculty of Medicine, Hasanuddin University, will evaluate whether the data are accurately collected. Since negligible risk to the patient is expected as both materials (MFAT and BCP) have been tested in other clinical trials [16,17,24], no data safety monitoring board will be installed. A safety report will be submitted every year to the Medical Research Ethics Committee, Faculty of Medicine, Hasanuddin University.
Medicine, Hasanuddin University. No interim analysis is deemed necessary.

**Amendments**

If deemed necessary, amendments to this protocol will be submitted to the ethical committee and competent authority and should be approved prior to implementation to ensure the safety and integrity of participants as well as the scientific value of the trial.

**Evaluation Methods**

**Safety Assessment Based on Physical Examination and Laboratory Measurements**

When an SAE occurs, it will be concluded that a combination of MFAT and BCP is not yet safe in the current setting. For AEs, if they do not occur at a higher frequency than that in patients treated with standard care (autologous bone) or can be resolved through noninvasive conventional methods (eg, analgesics or antibiotics), the combination of MFAT and BCP will be considered safe. In all other cases, combination of MFAT and BCP will not be considered safe yet.

**Radiographic Analysis**

To evaluate the success rate of the bone graft, the Bergland scale will be used [30]. This scale will evaluate the integrity and height of the alveolar bone graft and will classify bone height into 4 grades: grade I, bone height is almost normal; grade II, a bone height that is at least 75% of the interalveolar septum; grade III, a bone height of less than 75%; and grade IV, no evidence of bone integration [31].

**Histological and Histomorphometric Analysis**

Histological and histomorphometric analysis will be performed for at least 3 patients who received dental implants after alveolar cleft reconstruction, in accordance with previously published procedures [32]. Briefly, the implant preparation site will be developed using a trephine burr (2.0 mm × 10.0 mm in length) that allows biopsy specimen collection from the implant site without interfering with the regular procedure. The biopsy specimen will be fixed in 4% phosphate-buffered formaldehyde, dehydrated in an ascending series of ethanol, and embedded in 80% methylmethacrylate (BDH Chemicals) supplemented with 20% dibuthylphthalate (Merck); 8 g/L lucidol CH-50 L (Akzo Nobel), and 22 μL/10 mL N,N-dimethyl-p-toluidine (Merck). The biopsy specimens will be cut into 5-μm-thick sections and subjected to 2 different staining procedures (Goldner’s trichrome and Tartrate-resistant acid phosphatase staining). Several histomorphometric parameters (bone volume, osteoid volume, graft volume, and number of osteoclasts) will also be measured for quantitative analysis [32]. Two trained examiners will perform the histologic and histomorphometric analyses. In case of dispute, the biopsies will be reanalyzed to reach consensus.

**Statistical Analysis**

Since this is a single-arm safety study, statistical analyses will not be performed.

**Results**

The primary outcome parameter will be safety after 6 months’ follow-up, assessed by closely monitoring possible occurrences of AEs or SAEs, radiographic imaging, blood tests, and physical examinations. For efficacy, radiographic imaging will be used for clinical grading of the bone construct using the Bergland scale. In addition, bone parameters such as bone volume, ostoid volume, graft volume, and number of osteoclasts will be histomorphometrically quantified. We expect that the feasibility and safety of the procedure will be apparent, as well as its initial efficacy. Recruitment started in November 2019, and the trial is currently in the follow-up stage. This protocol’s current version is 1.0, dated September 15, 2019.

**Discussion**

In recent years, there has been increasing interest in the use of adipose tissue for cleft lip and palate reconstruction [33]. Its applicability mostly relies on the quantity of the tissue, the ease of surgical harvesting, and the type of surgical reconstruction in which the tissue is used, for example, correction of cleft lip volume asymmetry [34,35], improvement of velopharyngeal insufficiency after cleft lip and palate repair [36,37], or as an extra flap in cleft palate repair [38-41]. In this study, we will make use of the BFP for bone reconstruction. The BFP is a specialized adipose tissue rich in vascular supply, which is easy to harvest via the oral cavity during an intraoral surgery with minimal morbidity and discomfort [42].

Until now, there are only few reports on the use of adipose tissue as a regenerative compound for bony cleft reconstruction—a phase I clinical trial conducted by Khosravesh et al [24] and an animal study using ASCs for alveolar cleft repair [43]. Both studies used collagenase digestion of the tissue and culture expansion to obtain ASCs for personalized cleft reconstructions. An alternative is the SVF derived from adipose tissue via collagenase digestion, which requires a shorter time frame and may yield similar stem cell–like quantities, allowing intraoperative applications [44,45]. A previous clinical study by Prins et al [44] showed that addition of SVF in an intraoperative setting to calcium phosphate ceramics had an additive value on bone formation, implying that SVF can provide osteoinductivity when combined with calcium phosphate. However, so far, regulatory issues and relative expensive SVF production procedures prohibit its wide applicability [22,23]. Mechanically processed fat or MFAT has emerged as a rapid processing alternative to SVF and is being considered minimally manipulated and thereby less regulation restricted [22,23].

This is the first in human study evaluating a combination of MFAT and biphasic BCP as a regenerative graft for alveolar cleft reconstruction [46,47]. BCP is a ceramic scaffold with a balanced ratio between the less-soluble HA and the more-soluble TCP that results in mechanical and biological properties to support bone and cartilage tissue production [48]. It is sufficient for bone reconstruction in non–load-bearing applications and already accepted as standard of care for certain maxillofacial reconstructions [49].
Recently, calcium phosphate has been applied for alveolar cleft surgeries as well [16,17]. Patients within that study were treated at ages of 9-10 years, which is within the optimum age range for SABG [1]. However, we did not want to enroll children in a safety study with this novel concept in clinical practice. Therefore, although we realize that surgeries at a later age will (1) not make optimal use of the growth spurt and (2) may result in cases having larger or even critical size defects (which will not heal unless supplemented with grafts), we chose to only include older adolescent and adult patients, who are themselves capable of being involved in decision-making. We will perform this study in Indonesia because unoperated patients in this age group are difficult to find in Europe.

This is primarily a safety study, so the main conclusions of the study will be based on safety parameters, particularly on the occurrence of AEs or SAEs.

Acknowledgments

We would like to thank Nisrina Ekayani N, DDS, and Citra Dewi Sahrir, DDS, for helping in the administrative process of receiving ethical approval. DSNK received a scholarship from the Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Republic of Indonesia.

Data Availability

The data generated in the course of this study will be presented in the main manuscript reporting on the trial outcomes in a deidentified manner or provided in a multimedia appendix coupled with that manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.

References


Fatigue and Mental Illness Symptoms in Long COVID: Protocol for a Prospective Cohort Multicenter Observational Study

Ligia Pires¹, MD; Cláudia Reis², MD; Ana Rita Mesquita Facão², MD; Armin Moniri³,4,5,6, MD, PhD; Ana Marreiros⁴, PhD; Marta Drummond⁷, MD, PhD; Joana Berger-Estilita⁸,9,10, MD, PhD, MME

¹Serviço de Pneumologia, Hospital de Portimão, Centro Hospitalar Universitário do Algarve, Portimão, Portugal
²Serviço de Psiquiatria, Hospital de Portimão, Centro Hospitalar Universitário do Algarve, Portimão, Portugal
³Department of Otorhinolaryngology, Sahlgrenska University Hospital, University of Gothenburg, Gothenburg, Sweden
⁴Regenerative Medicine Program, Department of Biomedical Sciences and Medicine, University of Algarve, Faro, Portugal
⁵Algarve Biomedical Center, University of Algarve, Faro, Portugal
⁶Hospital Particular do Algarve, Faro, Portugal
⁷Centro de Responsabilidade Integrada Sono e VNI, Faculdade de Medicina, Universidade do Porto, Porto, Portugal
⁸Institute of Anaesthesiology and Intensive Care, Salem Spital, Hirslanden Medical Group, Bern, Switzerland
⁹Institute for Medical Education, University of Bern, Bern, Switzerland
¹⁰Centre for Health Technology and Services Research (CINTESIS), Faculty of Medicine, Universidade do Porto, Porto, Portugal

Corresponding Author:
Joana Berger-Estilita, MD, PhD, MME
Institute of Anaesthesiology and Intensive Care
Salem Spital
Hirslanden Medical Group
Schänzlistrasse 39
Bern, 3013
Switzerland
Phone: 41 788438161
Email: joanamberger@gmail.com

Abstract

Background: The aftermath of the COVID-19 pandemic continues to affect millions worldwide, resulting in persisting postvirus complaints and impacting peoples’ quality of life. Long COVID, characterized by lingering symptoms like fatigue and mental illness, can extend beyond a few months, necessitating further research to understand its implications.

Objective: This study aims to quantify the degree of physical and psychological fatigue in patients following COVID-19 infection and examine its correlation with mental health disorders.

Methods: Using a consecutive nonrandom sampling technique, we will conduct a prospective cohort multicenter observational study in 5 Portuguese hospitals. Symptomatic adult patients with previous COVID-19 attending follow-up consultations will be enrolled. We will include patients who had mild, moderate, and severe acute disease. We will assess clinical outcomes related to COVID-19, including the type of respiratory support such as high-flow nasal cannula, noninvasive ventilation, and invasive mechanical ventilation. The exclusion criteria will include previous severe psychiatric disorders confirmed by a psychiatrist; refusal or inability to respond to the questionnaire; concomitant neurological disorder; persistent fatigue symptoms during the 6 months before infection; and the need for invasive mechanical ventilation during COVID-19 infection due to a high prevalence of postintensive care syndrome. Our primary outcome is the prevalence of fatigue in patients with post–COVID-19 depression and/or anxiety, as measured by the Chalder Fatigue Scale (CFQ-11) and the Hospital Anxiety and Depression Scale (HADS). The secondary outcomes will include an assessment of health-related quality of life via the EQ-5D questionnaire and an exploration of the prevalence of symptoms of posttraumatic stress disorder (PTSD) using the 14-item Posttraumatic Stress Scale (PTSS-14). We will also examine the association between mental health symptoms and the severity of acute COVID-19. The post–COVID-19 data will be collected at least 6 months after the positive test and no longer than 9 months during the clinical appointment.

Results: We expect our multicenter study on patients post COVID-19 to reveal a significant link between mental illness symptoms and both physical and psychological fatigue. Patients with heightened depression and anxiety may report increased levels of...
fatigue. Additionally, we expect to find persistent PTSD symptoms in a subset of participants, indicating the enduring psychological impact of the virus.

**Conclusions:** This study may underscore the need for integrated care addressing physical and mental health in patients post COVID-19. The observed connections emphasize the importance of considering mental well-being for long-term health outcomes. Despite study limitations, our findings contribute valuable insights for future treatment strategies and highlight the necessity for comprehensive mental health support in post–COVID-19 care. This research provides valuable insights into the mental health implications of COVID-19 and its impact on post–COVID-19 fatigue and the overall well-being of affected individuals.

**Trial Registration:** ClinicalTrials.gov NCT05323318; https://clinicaltrials.gov/study/NCT05323318

**International Registered Report Identifier (IRRID):** DERR1-10.2196/51820

**KEYWORDS**
SARS-CoV-2; coronavirus; COVID; long COVID; fatigue; tired; tiredness; anxiety; depression; PTSD; stress; quality of life; mental health; post-COVID-condition; neuropsychological; neuropsychology; psychological; long COVID-19; COVID-19; myalgia; correlation; impairment

**Introduction**

Despite the World Health Organization (WHO) declaration marking the end of the COVID-19 pandemic on May 5, 2023 [1], millions of people continue to have postvirus complaints. Moreover, the virus remains endemic in many parts of the world. Confirmed COVID-19 cases have exceeded 430 million globally, with 200 million in Europe alone [2]. The causative agent of COVID-19 is the novel SARS-CoV-2 [3]. Although, as the name implies, respiratory symptoms are acute, the term long COVID (or post-COVID syndrome or long-haul COVID-19) began gaining recognition in the scientific and medical communities after the first descriptions of long-lasting symptoms related to mental health, such as anxiety or stress after the first infection [4]. While the definition of long COVID is unclear, the most frequent symptoms are fatigue and dyspnea [5,6]. Other less typical symptoms include cognitive and mental disorders, headache, myalgia, chest and joint pain, smell and taste dysfunction, cough, hair loss, insomnia, wheezing, rhinorrhea, sputum, and cardiac and gastrointestinal issues [4]. These symptoms may persist for up to 6 months after hospital discharge, severely impacting patients’ quality of life [7]. Since July 2021, long COVID is considered a disability under the Americans With Disabilities Act [8].

Per its definition, long COVID appears within 3 months after the onset of COVID-19, with symptoms lasting for at least 2 months that an alternative diagnosis [9] cannot explain, including, myalgic encephalomyelitis and chronic fatigue syndrome. However, studies have reported different persistent symptoms in contrasting durations and frequencies among survivors of long COVID [10,11]. Long COVID appears like other postviral syndromes observed in other coronavirus diseases. For example, symptoms of fatigue, myalgia, and psychiatric impairments have affected survivors of Middle East respiratory syndrome (MERS) and those with severe acute respiratory syndrome (SARS) for up to 4 years [10,12,13]. Even at 7-year and 15-year follow-ups, pulmonary and bone radiological complications were evident among a proportion of survivors of SARS who were predominantly younger than 40 years [10,14,15]. This is unsettling, as it implies that long COVID may extend beyond just a few months.

Fatigue, a primary persistent symptom of long COVID, has been reported in 10% to 70% of patients [16-19]. It is defined as “a decrease in physical or mental performance that results from changes in central, psychological, or peripheral factors due to the COVID-19 disease” [20]. Thus, post–COVID-19 fatigue depends on conditional and psychophysiological factors comprising the individual’s task, environment, physical, and mental capacity, as well as the disease’s central, psychological, and peripheral aspects [20].

While fatigue is not traditionally considered a neuropsychiatric disorder, it can be a symptom of many mental health conditions, such as depression, anxiety, and posttraumatic stress disorder (PTSD) [21,22]. Therefore, it is often included in discussions of mental health symptoms in post–COVID-19 syndrome [23]. The term “neuropsychiatric” often refers to a wide range of disorders affecting the brain and mental health. In contrast, “neuropsychological” refers to studying the relationship between brain function and behavior [24]. The primary focus of this study is on psychiatric symptoms, particularly depression and anxiety. However, we will also investigate PTSD, which is a neuropsychiatric disorder involving both neurological and psychiatric aspects. In addition, the health-related quality of life (HRQoL) assessment includes a section that evaluates aspects related to anxiety and depression. While not designed for diagnosing psychiatric disorders, this assessment can help gauge how these symptoms affect a person’s overall quality of life, making it more of a general health assessment.

The current literature on long COVID and its mental sequelae has relied on self-reported symptoms through questionnaires administered either in-person or through telephone interviews to comply with public health guidelines. However, the correlation between these self-reported symptoms needs to be clarified. There is an obvious need to accurately characterize the mental sequelae of COVID-19 and the risk factors associated with these outcomes.

This study aims to quantify the degree of physical and psychological fatigue in patients post COVID-19 and assess the correlation of fatigue with other mental sequelae—particularly depression, anxiety, and PTSD. Furthermore, we aim to explore its impact on HRQoL.
Our research questions for this study are outlined in Textbox 1.

**Textbox 1.** Research questions of the study.

<table>
<thead>
<tr>
<th>Research Question 1 (primary outcome):</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Do patients with long COVID who experience depression and/or anxiety symptoms have a higher prevalence of fatigue?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research questions 2 to 4 (secondary outcomes):</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is there a correlation between the type of fatigue and depression and/or anxiety symptoms among patients post COVID-19?</td>
</tr>
<tr>
<td>• Is there a correlation between fatigue and the presence of PTSD symptoms in patients post COVID-19?</td>
</tr>
<tr>
<td>• Do patients with fatigue associated with mental health disorders have a lower health-related quality of life (HRQoL) after COVID-19?</td>
</tr>
</tbody>
</table>

**Methods**

**Overview**

This protocol adheres to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statement [25]. This research is designed as an observational prospective cohort study.

**Recruitment**

Our study will be conducted in 1 private and 4 public Portuguese hospitals that have established follow-up post–COVID-19 medical consultations. Patients who meet the eligibility criteria will be invited to participate in the study at the end of their appointments in the follow-up clinics. This study will be conducted using a consecutive, nonrandom sampling technique.

**Inclusion Criteria**

Patients who cumulatively meet the following criteria will be eligible for inclusion: (1) ≥18 years; (2) previous case of COVID-19 at least 6 months after the diagnosis that is duly documented in the clinical record; (3) persistent symptoms after COVID-19 recovery, as defined by the WHO [9]; and (4) SARS-CoV-2 RNA confirmed by a positive real-time reverse-transcription polymerase chain reaction (RT-PCR) on a nasopharyngeal swab or SARS-CoV-2 antigen confirmed with a nasopharyngeal swab by a health care professional within 7 days of initial symptoms.

We will implement multiple measures to exclude the possibility of second infections and ensure that the individuals included have persistent symptoms following their initial COVID-19 recovery. Our inclusion criteria stipulate that individuals must have a positive test result for SARS-CoV-2 RNA or antigen within 7 days after their initial infection. Additionally, we will conduct thorough reviews of medical records and clinical assessments to assess symptom presentation and timing, differentiating between persistent cases and new infections. In cases of uncertainty or potential overlap with new infections, we are prepared to conduct repeat testing, which may include RT-PCR and genomic sequencing for accurate classification.

**Exclusion Criteria**

Patients who meet the following criteria will be excluded from participation: (1) preexisting psychiatric disorders diagnosed by a psychiatrist before contracting COVID-19; (2) inability to respond to the questionnaire; (3) concurrent neurological disorders, such as stroke with sequelae, Alzheimer disease, and Parkinson disease; (4) individuals subjected to invasive mechanical ventilation during their COVID-19 infection; or (5) those experiencing persistent fatigue symptoms within the 6 months before SARS-CoV-2 infection.

The study design and inclusion criteria will be reviewed regularly to ensure they align with current best practices and the latest understanding of COVID-19. A patient enrollment flowchart is presented in Figure 1.
**Data Collection**

Data collection will take place during appointments at the long COVID follow-up clinics in all participating hospitals. Qualified clinical study staff will gather the data using a clinical research form (CRF) (Multimedia Appendix 1). These individuals comprise health care professionals, such as nurses, clinical research coordinators, and health care experts who have received appropriate training and have the necessary expertise to conduct data collection and assessments in a clinical research setting. Their training ensures adherence to research protocols and ethical guidelines, guaranteeing accurate and consistent data gathering during clinical appointments. Two months after hospital discharge, we will mail all patients a letter offering a follow-up consultation in the outpatient clinic. Patients who agree to attend will be offered 2 appointments, consisting of an anamnesis and a physical examination, and given several self-administered tests in the waiting room. We will allocate 15 minutes for the questionnaire completion in the waiting room. In the first visit (T1, 3 months), the following data will be additionally collected: demographic characteristics like age, sex, and BMI (kg/m²); each participant’s medical history, such as the date of the first positive COVID-19 test (PCR or antigen) and smoking, alcohol, and drug consumption status; comorbidities and usual medication intake; and screening of symptoms and severity of the acute disease. Data collection in the second visit (T2, 6 months) will include post–COVID-19 symptoms, the Chalder Fatigue Scale [26] (CFQ-11), the Hospital Anxiety and Depression Scale (HADS) [27], the 14-item Posttraumatic Stress Scale (PTSS-14) adapted to COVID-19, and the EQ-5D [28] questionnaire.

**Ethical Considerations**

This study involves human participants and has adhered to all applicable ethical standards and procedures. The Algarve University Hospital Centre Ethical Committee approved this study (141/21; Multimedia Appendix 2). The patients’ race will not be included in the study because the ethics committee did not approve this differentiation. A paper-based consent form was previously approved by the ethics committee. This consent form has 2 components: information for the participant (regarding the project) and declaration of informed consent (to date and sign, in case of acceptance). The informed consent procedure, mailed to potential participants, will provide them with detailed information about the study’s objectives, risks and benefits of participation, and data management practices (Multimedia Appendix 3). The study will follow ethical guidelines and regulations related to informed consent, and participants will have the opportunity to ask questions and provide voluntary consent before enrolling. Participants will be given a month for reflection before written informed consent. As for data management, the study will implement a secure and confidential system that uses data encryption and secure storage procedures to protect participant confidentiality and comply with data protection laws and regulations. Only authorized individuals who have a legitimate need to access the data will be permitted to do so, and any data sharing will be done in compliance with the relevant ethical and legal requirements. As applicable, all procedures from this investigation followed the Declaration of Helsinki. All researchers will comply with the Data Protection Acts of their respective academic institutions.

**Primary Outcome**

**Fatigue**

A validated fatigue assessment tool will be used at the follow-up visit to capture a broader range of participants. This will provide a more comprehensive understanding of the relationship between fatigue and mental symptoms in individuals with long COVID. Fatigue will be measured using the CFQ-11 [26]. This is an
11-item self-report measure of physical and mental fatigue. Participants rate their fatigue experienced over the past month compared to their usual energy levels using a 4-point Likert scale ranging from 0 (“less than usual”) to 3 (“much more than usual”) [29]. Scores range from 0 to 33. A bimodal version can also be calculated, where scores range from 0 to 11, with a cutoff score of 4 or more indicating a case of fatigue [29]. Reliability coefficients for the CFQ-11 have been shown to range between 0.8 and 0.9 [29-31] in both patients with chronic fatigue and the general population. This study will use only the Likert scale to reduce the risk of ceiling effect bias.

Secondary Outcomes

Depression and Anxiety

Symptoms of depression and anxiety will be assessed using the HADS [27], a practical, validated tool for assessing symptom severity in anxiety and somatic disorders in psychiatric and primary care settings and the general population [32]. This measure comprises 14 items, 7 measuring symptoms of anxiety and 7 measuring symptoms of depression. The scores are categorized into different ranges: 0 to 7 as “normal,” 8 to 10 as “mild,” 11 to 14 as “moderate,” and 15 to 21 as “severe.” The HADS questionnaire has been validated in many different languages and settings [32-34].

HRQoL Assessment

HRQoL will be assessed using the original EQ-5D questionnaire, which comprises 2 parts. The first part is the EQ-5D-3L questionnaire [28], a health state classification scheme of 5 items (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), each having 3 alternatives (1=no problems, 2=moderate problems, and 3=severe problems). The combination of answers for the 5 items represents the health state (given as an index), ranging from 0 (worst possible health state) to 1 (best possible health state). The second part comprises the EuroQol Visual Analog Scale (EQ-VAS), which represents health states in a range from 0 (the worst possible health state) to 100 (the best possible health state). The EQ-5D can distinguish between the health conditions of patients with diverse injuries [35,36] and has been validated in several populations [37].

PTSD Assessment

This study includes an assessment of PTSD as a secondary measure. It aims to evaluate the potential impact of PTSD on the primary outcome of fatigue and provide a better characterization of the study population. Although previous PTSD to COVID-19 infection is an exclusion criterion, some participants may still have experienced PTSD symptoms after their COVID-19 infection, so evaluating the prevalence and severity of these symptoms in the study population could be relevant.

Given that the “gold standard” clinician-administered, semistructured psychiatric interviews to diagnose PTSD symptoms were not feasible during the workload of the COVID-19 follow-up clinic, we chose the patient-reported outcome instrument PTSS-14. This scale was developed based on the Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (DSM-III-R) criteria for PTSD and subsequently validated for patients following intensive care [38]. Patients were asked to assess the frequency of 10 common PTSD symptoms on a 7-point Likert scale from 1 (never) to 7 (always). Item scores were summed up to a total score (ranging from 10 to 70). A total score above 35 was suggested to indicate clinically relevant PTSD symptoms. A similar 14-item version of PTSS was later developed to reflect the changes in the fourth edition of the DSM (DSM-IV), where a cutoff value of 45 indicated presumable PTSD [39]. We modified the latter to include COVID-19 patients—whenever the term “intensive care unit (ICU)” appeared, we substituted it with “COVID-19.” We applied the same diagnostic cutoff to this adapted scale.

Association of Mental Illness Symptoms With the Severity of Acute COVID-19

COVID-19 severity will be categorized according to the WHO classification [40]. The classification includes the following: mild illness (mild symptoms without the radiographic appearance of pneumonia); pneumonia (having symptoms and the radiographic evidence of pneumonia, with no requirement for supplemental oxygen); severe pneumonia (having pneumonia, including one of the following: respiratory rate >30 breaths/minute, severe respiratory distress, or blood oxygen saturation level ≤93% while at rest and breathing room air); and critical cases (eg, respiratory failure requiring invasive mechanical ventilation or nasal high flow oxygen, septic shock, other organ failure occurrence, or admission to the ICU).

Statistical Analyses

The statistical analysis methodology is designed to align with the study’s defined objectives. All analyses will be conducted using SPSS (version 27; IBM Corp) and R (R Foundation for Statistical Computing) software. The analyses will be carried out in 5 phases, as outlined in Textbox 2.
Textbox 2. The 5 phases of the statistical analyses carried out in this study.

- **Phase 1: Characterization of the patient sample**
  - 1.1. Sociodemographic analysis of the variables collected at T1 and the variables from the surveyed instruments at T2. This involves descriptive analysis, both univariate and bivariate, which will be differentiated by:
    - 1.1.1. Quantitative variables: mean and SD, coefficient of variation, and median and IQR
    - 1.1.2. Qualitative variables (nominal and ordinal categorical): analysis of distributions through absolute and relative frequencies, including absolute and relative accumulated frequencies
    - 1.1.3. Quantitative variables will be reported alongside appropriate tests for normal distribution to ensure the accuracy and validity of our statistical analyses.

- 1.2. Bivariate and descriptive analysis using 2D tables will be calculated for all qualitative variables belonging to T1 with all instrument scores collected at T2: Chalder Fatigue Scale (CFQ-11), Hospital Anxiety and Depression Scale (HADS), 14-item Posttraumatic Stress Scale (PTSS-14) adapted to COVID-19, and EQ-5D plus recorded COVID-19 symptoms. The previous points will be reiterated in this bivariate context, utilizing the computational aid of split-file categorical variables in situations where one variable is metric and the other is categorical.

- 1.3. Internal consistency and reliability analysis of the T2 instruments: calculation of the Cronbach alpha (general and partial) for all ordinal items and item subgroups differentiated by the subscales defined by the respective authors

- **Phase 2: Determination of fatigue incidence caused by long COVID and its relationship with T1 and T2 variables**
  - 2.1. Procedures for determining fatigue cutoff points in this cohort of patients:
    - 2.1.1. Bivariate analysis using optimal binning procedures for categorical variables (age class, sex, and BMI) with the total fatigue score, as well as partial scores of physical and psychological fatigue
    - 2.1.2. Multivariate analysis by constructing models in the format of decision trees via the CHAID (chi-square automatic interaction detection) algorithm [41], with total and partial fatigue scores as dependent variables

- 2.2. Inferential analysis based on variables resulting from 2.1.1 and/or 2.1.2, considering the complexity of variables in T1 and T2:
  - 2.2.1. (1) Chi-Square tests to capture the association of categorical variables; (2) Shapiro-Wilk tests to verify the distribution of quantitative variables; (3) tests for comparing 2 or more population means: parametric analysis of variance (ANOVA) test and nonparametric (Kruskal-Wallis) test; and (4) correlation tests (Spearman or Pearson) to verify the degree of association between T2 instrument scales

- **Phase 3: Determination of patient groups based on fatigue differentiation**
  - 3.1. Conducting classification analysis using clustering techniques, specifically the 2-step cluster analysis, to explore variables with greater power in differentiating groups of patients

- **Phase 4: Exploration of the best combination of T1 and T2 variables capable of discriminating patient fatigue levels through a discriminant multivariate analysis**

- **Phase 5: Investigation of the association between phase 3 groups and all T2 instrument scales using chi-square association tests**

### Power and Sample Size

We plan to conduct a prospective multicentric observational study to investigate the prevalence and severity of mental illness symptoms and fatigue in individuals with long COVID. Based on previous studies and expert opinion, we expect the prevalence of fatigue and mental symptoms in this population to be around 50% [42,43]. We aim to detect a minimum effect size of 0.2 with a power of 80% and a significance level of \( P=0.05 \).

Using these assumptions, we performed a sample size power analysis using a 2-sample \( t \) test. We found that a total sample size of 200 participants (100 with adverse mental health symptoms and 100 without) would be required to detect the minimum effect size of 0.2 with 80% power and a significance level of \( P=0.05 \). The required sample size was calculated using an a priori power analysis with an online calculator [44]. To compensate for a projected 25% loss to follow-up, we aimed for 250 participants.

We plan to recruit participants from multiple hospitals and primary care centers across Portugal, focusing on individuals who have been diagnosed with long COVID and are experiencing mental illness symptoms and/or fatigue. Recruitment will be closely monitored throughout the study to assess the pace at which participants are enrolled. The need for sample size adjustment will be determined based on the observed recruitment rates and the accumulating data. We plan to conduct an interim analysis after the recruitment of the first 100 patients.

### Results

This study received no specific grant from any funding agency in public, commercial, or not-for-profit sectors. Recruitment began in October 2021 at Portimão Hospital and Alvor Hospital.
in June 2022 at Faro Hospital and São Sebastião Hospital, and in October 2022 at Fernando Fonseca Hospital. The deadline for the end of the recruitment period in all centers was December 2023. The preliminary study results will be published in a peer-reviewed international medical journal after October 2023.

The study timeline is shown in Figure 2. This study is included as part of the first author’s (LP) PhD thesis. The initial study results were presented at the PhD interview in November 2023. We aim to publish the study in an indexed scientific journal and present the results at national and international congresses.

We plan to enroll 250 participants, including 100 with adverse mental health symptoms and 100 without, to explore the intricate relationship between adverse mental health symptoms and both physical and psychological fatigue in individuals recovering from long COVID. Preliminary analyses indicate a compelling connection, wherein patients who report heightened levels of depression and anxiety also tend to experience increased fatigue. These findings underscore the importance of considering mental health as a pivotal factor in understanding the enduring impact of COVID-19 beyond the acute phase. In addition to depression and anxiety, our study explores the persistence of PTSD symptoms in a subset of participants. Initial results suggest that, even after recovery from the acute phase of the infection, a proportion of individuals continue to grapple with the psychological repercussions of the virus. This observation aligns with emerging evidence suggesting a prolonged psychological impact of COVID-19, emphasizing the need for comprehensive mental health support for patients after contracting COVID-19.

The identified links between mental health symptoms and fatigue have broad implications for the holistic care of this patient population. Understanding these connections can guide targeted interventions, emphasizing the importance of addressing mental well-being alongside physical recovery. As we delve deeper into the data and conduct further analyses, a more nuanced understanding of these relationships will emerge, informing future health care strategies and interventions.

**Discussion**

**Expected Findings**

As the COVID-19 pandemic ceases, more patients enter the chronic phase of the disease. Identifying groups at high risk of cognitive and psychiatric dysfunction may allow for targeted intervention to effectively meet their physical, neurological, and psychological health care needs. Our study, conducted in the aftermath of the COVID-19 pandemic, provides a unique opportunity to investigate the persisting health effects of the virus on those who have recovered from the acute phase of the disease. We specifically aim to explore post–COVID-19 fatigue and mental health disorders using specific tools to separate physical and psychological fatigue symptoms. We also use a prospective cohort multicenter study design, which is less prone to bias, and we aim to include both hospitalized and nonhospitalized patients, covering all degrees of acute COVID-19 severity. We anticipate that our research will reveal a significant association between the presence and severity of mental health symptoms, including depression, anxiety, and PTSD, and the degree of physical and psychological fatigue in patients following COVID-19. We expect to observe that patients experiencing more pronounced mental symptoms will report higher levels of physical and psychological fatigue.

The potential links between mental symptoms and fatigue suggest a connection between the mental well-being of patients post COVID-19 and their experience of fatigue, encompassing both physical and psychological dimensions. The existing research has indicated a high prevalence of anxiety and depression symptoms among patients hospitalized with COVID-19 infection, with a quarter of patients experiencing at least mild symptoms of acute stress disorder [42]. Moreover, a study conducted in Iran after the outbreak of COVID-19 found a correlation between the prevalence of chronic fatigue syndrome and PTSD [45]. The study reported that 5.8% of subjects experienced PTSD symptoms 6 months after the onset of SARS-CoV-2 infection. Interestingly, the study also noted that female sex was associated with a higher risk of fatigue. At the same time, variables such as oxygen saturation at admission, primary symptoms, ICU admission, and laboratory test parameters did not show a significant association with fatigue occurrence.

Based on these findings, we anticipate that patients post COVID-19 may exhibit more pronounced symptoms of...
depression and anxiety and heightened levels of physical and psychological fatigue. This insight emphasizes the importance of considering mental health as a crucial factor in understanding the long-term effects of COVID-19. It underscores that the impact of the virus extends beyond the acute illness phase and can manifest in persistent mental symptoms that influence patients’ overall well-being. Moreover, it can imply that addressing anxiety symptoms in post–COVID-19 care may have a positive impact on reducing fatigue and improving patients’ quality of life. We also aim to explore the prevalence and severity of PTSD symptoms in a subset of patients following COVID-19. While this is not the primary focus of our research, assessing the presence of PTSD symptoms in this study population allows for a better characterization of the patient group. Some participants continue to experience PTSD symptoms following COVID-19 infection, suggesting that the psychological impact of the virus can be long-lasting. These findings may shed light on the need for comprehensive mental health support for patients post COVID-19.

The implications of our potential findings are widespread. They underscore the importance of integrated care for patients post COVID-19 that addresses both physical and mental health aspects. Clinicians and health care providers will then be aware of the potential for persistent mental symptoms and their impact on patients’ fatigue levels and overall HRQoL.

Limitations
Our study’s limitations include the possibility of sampling bias, as individuals with more severe symptoms of fatigue or other mental health symptoms may be less likely to participate. This possibility could result in underrepresentation and bias within the study’s findings. To minimize this, we will use diverse recruitment methods across multiple sites and collect data on participants’ reasons for declining participation. We will also provide detailed information about the study and its potential benefits to participants while ensuring an ethical and respectful recruitment process sensitive to participants’ health concerns and symptoms.

We also acknowledge the limitations related to our exclusion criteria and use of the CFQ-11, which has yet to undergo specific validation for patients with COVID-19. This recognition underscores the importance of future research endeavors to validate the applicability of the CFQ-11 within this patient population, despite its utilization in prior studies [46–48].

While there are valid concerns regarding the exclusion criteria and assessment of psychiatric disorders in our study—including the fact that assessing preexisting fatigue may be difficult—we have devised a plan to mitigate these concerns. We intend to meticulously collect detailed medical histories and review the medical records of all enrolled patients to ensure the precise application of the exclusion criteria. However, a notable limitation of our study arises from the absence of a baseline assessment conducted before the onset of COVID-19 infection. This absence impedes our ability to comprehensively evaluate the influence of COVID-19 on various health parameters, such as fatigue and mental health. The absence of a baseline assessment may also limit our capacity to distinguish between preexisting conditions and post–COVID-19 effects, which is a key consideration in understanding the full spectrum of the virus’s impact on individuals. Additionally, while we acknowledge that the measurements used in the study are based on self-report—and therefore subjective—we emphasize that self-reported questionnaires are commonly used in scientific studies. These questionnaires hold significance in assessing the extent of physical and psychological fatigue in patients following COVID-19. Despite these limitations, our study is poised to make a significant contribution to the existing literature by shedding light on the impact of COVID-19 on mental health. This research has the potential to guide future treatment and management strategies, serving as a valuable resource for the health care community.

Conclusion
This study protocol outlines a critical investigation into the lingering physical and psychological effects of long COVID. It emphasizes the importance of understanding the ongoing health challenges individuals face even after recovering from the acute phase of the virus. This study’s research questions focus on assessing the potential correlations of mental illness symptoms (ie, physical and psychological fatigue) among patients post COVID-19.

In the aftermath of a global pandemic, this research is timely and crucial. It has the potential to inform health care strategies and interventions that provide targeted and holistic care to individuals grappling with long COVID. Ultimately, this study aims to improve patient outcomes, enhance the quality of life for those affected, and contribute to broader efforts to address the multifaceted health implications of this unprecedented global crisis.

Authors' Contributions
LP conceptualized the study, wrote the first draft of the protocol, and planned the study. CR and RF helped write the first draft of the protocol. AB gave critical input throughout the study.

JMIR Res Protoc 2024 | vol. 13 | e51820 | p.234 https://www.researchprotocols.org/2024/1/e51820

JB-E wrote the first draft of the protocol and gave critical input throughout the study. AM was responsible for the statistical planning and methodology. MD planned the study. All authors contributed to the manuscript and approved the submitted version.

No generative artificial intelligence (AI) was used in any portion of the manuscript writing.

Conflicts of Interest
JB-E is an associate editor for BioMed Central (BMC) Medical Education and has received travel expenses from Medtronic for the “Save the Brain Initiative” training. The remaining authors have no conflicts of interest to disclose.
References


44. Sample size calculators for designing clinical research. Sample Size Calculators. URL: https://sample-size.net/ [accessed 2023-12-30]


Abbreviations

CFQ-11: Chalder Fatigue Scale
CRF: clinical research form
DSM-III-R: Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
EQ-VAS: EuroQol Visual Analog Scale
HADS: Hospital Anxiety and Depression Scale
HRQoL: health-related quality of life
ICU: intensive care unit
MERS: Middle East respiratory syndrome
PTSD: posttraumatic stress disorder
PTSS-14: 14-item Posttraumatic Stress Scale
RT-PCR: reverse-transcription polymerase chain reaction
SARS: severe acute respiratory syndrome
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
WHO: World Health Organization

©Ligia Pires, Cláudia Reis, Ana Rita Mesquita Facão, Armin Moniri, Ana Marreiros, Marta Drummond, Joana Berger-Estilita. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.01.2024. 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 Impact of Forced Separations Between Women and Their Pets in Domestic Violence Situations and the Effectiveness of Crisis Response: Protocol for a Conceptual Framework

Jasmine Montgomery¹, BSW, MCouns; Janice Lloyd¹, DCR, BSc (Hons), Grad Cert Edu, MEd, PhD; Zhanming Liang¹, Cert SQIL, MBBS, MSc, PhD

College of Public Health, Medical and Veterinary Sciences, James Cook University, Queensland, Australia

Corresponding Author:
Jasmine Montgomery, BSW, MCouns
College of Public Health, Medical and Veterinary Sciences
James Cook University
1 James Cook Drive
Queensland, 4811
Australia
Phone: 61 47814731 ext 07
Email: jasmine.montgomery@my.jcu.edu.au

Abstract

Background: Women are at high risk of experiencing trauma, guilt, and stress when forced to separate from their companion animals when fleeing domestic violence. Where little support is available for women and pets to stay together, women may be forced to delay leaving the abusive relationship or leave the pet with the abuser. Forced separation places both women and pets at substantial risk, where pets may be used as a coercive control measure. However, little evidence exists regarding the extent to which Australian services or policies offer support in these circumstances.

Objective: This research aims to increase the understanding and the impacts of forced separation between women and their pets in domestic violence situations. The research will investigate the effectiveness of service responses for both women and animals, aiming to develop a policy framework that guides service improvement with the goal of enhancing outcomes for women and pets fleeing domestic violence.

Methods: This protocol paper describes the process of developing a conceptual framework of 4 studies that include a scoping review, policy analysis, focus groups, and interviews that guide the design of the qualitative research project.

Results: A scoping review of the literature on forced separation from pets in domestic violence, natural disasters, and homelessness situations has led to the development of a conceptual framework that guided the design of the proposed study. The review also confirmed the necessity of the proposed research project in addressing the lack of Australian national frameworks and guidance available for women and pets seeking formal support in domestic violence situations. As of August 2023, supporting organizations have commenced the distribution of the research flyers. Expected data collection will be completed between August and October 2023. The results are expected to be published in June 2025.

Conclusions: Via a systematic process, the importance of the proposed study in improving the understanding of the impact of forced separation between women and their pets at times of domestic violence and the gaps in best supporting both women and their pets has been confirmed. A study design based on the learnings from previous studies and the focus of the current research has been finalized. The impact of the research project in developing an Australian national framework for best supporting women and their pets in crisis situations is anticipated.

International Registered Report Identifier (IRRID): PRR1-10.2196/52067

(JMIR Res Protoc 2024;13:e52067) doi:10.2196/52067

KEYWORDS
companion animal; domestic violence; forced separation; research protocol; animal welfare; pets; animal abuse; Australia; coercive control; victim; abusive partner; abusive; women; trauma; support; animal
Introduction

In Australia, 69% of households live with a companion animal (pet), and 86% of households with a pet have children [1]. Dogs and cats are the most popular type of pets [1]. The main reason for living with a companion animal is companionship [1]. The relationship is considered beneficial both psychologically and physiologically for humans and animals [2]. Pets are a vital support system providing emotional support or strength at times of domestic violence [3]. Survivors in domestic and family violence situations often live in terror and face threats to themselves and their pets [4]. Sadly, women in domestic violence situations are often faced with the torturous decision to leave their pet with the perpetrator to seek safety or access temporary fostering, resulting in forced separation from their companion animal (Montgomery et al [5], in press), thus losing the emotional support normally received from the relationship [3].

Barrett et al [6] found that decisions to leave or stay in the relationship were impacted by the concerns for the animal’s welfare, with 56% of women delaying leaving the relationship to protect their pet. Women with both children and pets were also found to delay leaving an abusive relationship out of concern for the pet’s welfare [7]. Most women who delayed were forced to leave their companion animal with the perpetrator when they eventually fled to safety and 47% of women would have fled to safety with their companion animal if support was available [8]. Completing a safety plan when leaving domestic violence situations was often compromised due to a lack of pet-inclusive shelters, often leading to homelessness in order to stay with their pet [7]. When survivors are forced to leave their companion animals with the perpetrator, the risk of coercive control (such as monitoring a person’s movements) increases where the companion animal is used as a coercive control tool [9]. The companion animal in this situation may be subject to continued maltreatment [9], often resulting in torture or death [10] and survivors experience additional guilt and trauma [3] as a result. Often, they consider returning to their partner for the sake of their companion animals’ safety [8]. Where companion animals have survived domestic violence, signs of distress in the animal have been observed through behavioral changes, such as avoidance and vocalization [4,11]. Devastatingly, in Australia, such behavioral changes often result in euthanasia of the pet [4,11].

The emotional attachment between survivors of domestic violence and their pets may be substantial due to sharing the experience of abuse [4], which makes a deliberate act of cruelty or death of a companion animal particularly torturous [12]. While it is the case that domestic violence is a human issue that affects both men and women, it is recognized as a gender-based issue where men are more likely to perpetrate violence against women and is considered an epidemic problem that requires change in Australia [13]. A recent report on homicide in Australia [14] reveals that, from 1989 to 2020, the incidence of intimate partner homicide is consistently much higher for female survivors than male survivors. The most recent statistic (2019–2020) states that female individuals were the targets in 36 (80%) of the 45 intimate partner homicides. Considering Australia is one of the highest pet ownership countries in the world, where women with children are more likely to have a pet [15], it is vital to address the risks for survivors and their companion animals at times of forced separation because of domestic violence. In such a context, a research project has been developed to investigate the existing policy framework and relevant services that provide support to people and companion animals in domestic violence situations. This protocol paper will explain the process of confirming research gaps and determining research questions and will provide details of the overall project design to be used by the proposed project as informed by the learnings from previously published studies.

Methods

Overview

A scoping review [5] using the keywords “human-animal relationship/bond,” “pets,” “companion animals,” “animal abuse,” “violence,” “homelessness,” “housing,” and “disasters,” was conducted between March and August 2022. The review focused on identifying empirical studies on the human-animal relationship and crisis or situational change with no date limitation. The review was guided by Arksey and O’Malley’s [16] framework for scoping reviews and conformed to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist for scoping reviews [17]. English-language, scholarly peer-reviewed papers that included adults with a strong relationship with a pet and an event or change of situation were a criterion for the scoping review. All methodology types were accepted. Gray literature and certain animal types (rodents, wildlife, zoo animals, and working animals) were excluded. The papers were assessed on their ability to fit within the inclusion criteria. Five databases (MEDLINE Ovid, PsycINFO, Scopus, CINAHL, and Emcare Ovid) were searched, and a total of 42 scholarly papers that met the inclusion criteria were identified and included for data extraction. The scoping review mapped the concept of forced separation between people and their companion animals in areas of crisis or situational change and examined policies that included companion animals. The study design and methods used for the studies were also examined to inform the current project design. Please see the full list of papers included in the scoping review in the Multimedia Appendix 1 [3,4,6-11,18-51].

Scoping Review Findings That Informed the Development of the Protocol

The identified studies in [5] scoping review were predominantly quantitative and conducted in the United States, with a focus on the co-occurrence of animal abuse and domestic violence. The lens of research has recently focused on the relationship and animal maltreatment or welfare concerns. Surveys and semistructured interviews were the common forms for collecting quantitative studies and qualitative data respectively. The average sample size consisted of 200 participants for quantitative studies and 20 participants for qualitative studies. The target population was predominantly female adults seeking refuge from domestic violence shelters and support services.

https://www.researchprotocols.org/2024/1/e52067

JMIR Res Protoc 2024 | vol. 13 | e52067 | p.239

(page number not for citation purposes)
The scoping review [5] confirmed a lack of support for both humans and animals at times of forced separation because of domestic violence. The oversights of the animals’ safety and welfare showed that animals were being left with the abuser [18] and women delayed leaving a violent relationship to protect the companion animal [19]. Additional barriers that were identified included geographical locations, lack of available supports [19], lack of awareness of supports, and attachment or fear of separation from the companion animal [4]. As a result of these barriers, the risks to safety, health, and well-being for women, children, and their companion animals have increased.

The scoping review [5] findings revealed survivors were often reluctant to reach out to services due to a lack of trust in accessing support services, veterinary care, and law enforcement. A lack of trust was associated with a fear of being forced to separate from their companion animal [10,18,20,21]. The reluctance to access support, and the responsibility weighing on women to access supports [6,22] is highly concerning. Although many studies in the literature provided implications for service providers, no research was found that investigated the policy frameworks that provide support to people and companion animals in domestic violence situations at any system, organization, societal, or individual level [5].

Ethical Considerations

The following ethical considerations are guided by the Global Women’s Institute for the Department of Foreign Affairs and Fair Trade [52], which provides recommendations for projects specific to researching women in domestic violence situations. Ethical approval was granted from the Human Research Ethics Committee (approval H9148). Participation in this study is voluntary and written and verbal consent will be obtained from every participant. The data will be retained for a minimum period of 5 years and will only be accessible to the research team. All data collected will be deidentified and pseudonyms will be provided. All audio recordings for both target populations will be erased after transcription. The primary target population will have the opportunity to review the transcriptions in writing via email. To avoid comprising anonymity and confidentiality for the primary target population, specific locations, age, occupation, culture, and religious discourse in the primary target will not be included in the narrative where there is potential to make the participant identifiable. Consent will be obtained verbally prior to the commencement of the focus group discussion and interviews. Participants are reminded of the voluntary nature of the study and their rights to not answer questions or withdraw their participation from the study. The focus group will be informed that confidentiality is not guaranteed and will be requested to anonymize discussions of their opinions and keep the group discussions private. Confidentiality and anonymity are provided to the interview participants.

Research Focus and Research Questions

The research aims to inform the Australian policy framework by investigating how support services operate across different contexts for adult women and their companion animals affected by forced separation to reduce negative impacts for both people and animals when fleeing domestic violence situations. The research aim will be achieved by the following two objectives:

1. Identifying the impacts of forced separation between adult female survivors of domestic violence and their companion animal’s health, safety, and living conditions.
2. Identifying the existing strategies and support services, the perceived effectiveness of these strategies, and areas for improvement to develop recommendations that maximize support to people and their companion animals fleeing domestic violence situations.

The research seeks to answer the following two questions.

1. How does forced separation impact the domestic violence survivor and their companion animal under the existing policy and support framework in Australia?
2. What are the factors and how do these factors influence the extent that the benefits of the existing services currently available to people and their companion animals are realized?

A qualitative design will be used to address the gaps in the literature of a lack of national framework to guide pets and women in domestic violence; the impact of forced separation; and the roles, attitudes, and beliefs of seeking and providing services to better understand the impacts and perceptions of forced separation. The transformative paradigm views privilege and power as a social construction that is embedded through social, political, cultural, economic, gender, age, disability, race, and ethnicity. The transformative worldview is a suitable framework providing the lens of power and oppression with a focus on positive social change [53].

Conceptual Framework

Based on the findings of the scoping review [5] and the role of support services in preventing or minimizing adverse outcomes due to forced separation, a conceptual framework (Figure 1) was developed. The framework indicates that policy and adequate, effective support services are required to improve the outcomes for people and companion animals who must leave their homes because of domestic violence. The scoping review [5] confirms that a policy framework, key supports, and elements required to achieve these outcomes remain unclear. It is important to understand existing policies, support services or providers, and those who use the services in Australia so that improvements can be made to best support people and companion animals fleeing domestic violence. Guided by the conceptual framework created for this study, 4 steps (Figure 1, studies 1–4) need to be implemented to enrich our understanding of the key elements leading to the development of a policy framework on the forced separation of companion animals because of domestic violence that is relevant to the Australian context. The steps include:

2. A policy or services analysis and a scanning of the key supports to humans and animals that will analyze the purpose; construction; implementation; and impacts to understand, evaluate, and provide meaning and context [54].
3. Semistructured individual interviews with participants who have accessed a variety of support services (refuges, crisis services, animal welfare services, and mainstream such as women’s legal services) will be conducted to increase the understanding and impacts of forced separation on people and companion animals.

4. Focus groups with staff and service providers will be conducted to understand and identify perceptions of the effectiveness and adequacy of service provision.

A critical analysis of steps 2-4 will be completed to compare the findings of the most common types of support, service gaps, and availability of services, leading to the development of an improved policy and support framework.

Figure 1. The conceptual framework. RQ: research question.

**Project Design and Method**

**Participants: Characteristics and Recruitment**

The research project has 2 target populations. The primary target population is female individuals who have or have had considered themselves to have a strong emotional bond in a relationship with their companion animals; have experienced domestic violence; and have been forced to separate from their companion animals or sacrificed their own health, safety, and living arrangements to stay with the animals. The secondary target population is those who have not been directly affected by forced separation and domestic violence but have provided or are currently involved in providing professional support services to the primary target population. Due to the complex and sensitive phenomenon, both target populations are adults, 18 years and older [52].

Both target populations will be recruited through relevant domestic violence or animal welfare organizations. The organizations will be responsible for making direct contact with the potential participants via emails or organization-based advertisements. Potential participants will be encouraged to make direct contact with the principal researcher (first author) should they wish to participate in the study. Both target populations will be geographically recruited nationally across all states and territories in Australia. Due to financial and logistic constraints [55], both target groups are required to be fluent in the English language.

**Sampling Strategy**

Purposive sampling will be used to gather specific characteristics of survivors who have a strong emotional bond with their companion animals (primary target population) to maximize the richness of the data in addressing the research questions [56]. People who received an invitation from their perspective organization and made contact to participate in the study will need to fit the characteristics of either target population 1 or 2.

**Key Stakeholders**

To maximize the ethical sensitivity of the research, an advisory group of professionals in the field of domestic violence, advocacy bodies, and animal welfare organizations will be created [52]. Experts participating in the advisory group are...
excluded from the secondary target population. The principal researcher candidate and 2 research advisors will hold meetings via Zoom (Zoom Video Communications) prior to ethics submission and after analysis. The advisory group is sought for their expertise regarding sensitivity, recruitment pathways, research questions, and participation sheets. After the analysis, for advice on recommendations based on the findings from the study.

Data Collection Tools

In-depth individual interviews are best suited for “sensitive” populations [52] and web-based options may increase the participant response rate [57]. Hence, the primary target population will be invited to participate in individual semi-structured, web-based interviews via Zoom. These interviews are expected to take around 1 hour and will be audio recorded. Focus groups are well suited to discussing beliefs, opinions, and attitudes surrounding programs [52], interventions, and service gaps [58]. Therefore, the secondary target population will be invited to participate in web-based focus groups with audio recordings via Zoom. There will be 4 focus groups nationwide. The focus groups are estimated to last 1 to 2 hours as it is important to allow time during the focus groups for rapport building and voicing opinions [56]. Both types of interviews will be professionally transcribed. Verbal and written consent to participate will be obtained from all participants. All participants will be given the opportunity to review a summary of the transcriptions prior to publication [59].

Sample Size

The average number of participants in related qualitative studies identified in the scoping review was 20 (Montgomery et al [5]). The method of the research project is designed to gather in-depth, rich data or high-quality dialogue [60]. Hence, between 12 and 20 participants will be sampled from the primary target group, with the final number of participants being guided by data saturation of main themes, and no new insights or issues are found [61]. The secondary target population will consist of 4 focus groups throughout Australia. When a group consists of high knowledge, a minimum of 4 participants are required to develop accurate information [62] and the probability of identifying themes with 6 participants is higher than 99% [63]. Due to the expertise and knowledge of the participants, there will be a minimum of 4 and a maximum of 8 participants to allow for space and reflection with each group member [56]. The number of participants for the target populations is supported by a recent systematic review of effective sample sizes for saturation in qualitative research [61].

Data Analysis

Interpretive work is required to identify meanings and themes from participants’ opinions, perceptions, and experiences to meet the research aims and overall purpose. Thematic analysis will be used to provide a systematic approach to coding and conceptualizing themes [58]. Areas of analysis will include the impacts and outcomes of forced separation, accessibility of services, types of unmet needs, experiences of accessing services, and benefits of existing services. When the analysis of each step is completed, a critical analysis will be completed to aggregate the data [58] to provide a complete picture of the policy framework [54]. NVivo 12 software (Lumivero) will be used to facilitate the data analysis process.

Results

A scoping review of forced separation of companion animals in crisis situations has been completed, identifying the research gaps and guiding the research questions and design for the research project. As of August 2023, supporting organizations have commenced the distribution of the research flyers. Expected data collection will be completed between August and October 2023. The results are expected to be published in June 2025.

Discussion

Expected Findings

It is expected that the findings will identify the substantial issues experienced by women and pets in domestic violence situations such as psychological distress, grief, loss, and the complexity of decision-making when considering a pet. It is expected that women and pets need to be considered more seriously in Australia and the development of policies and services needs to include the consideration of pets in safety planning, accommodation, and long-term housing as their standard practice.

Comparisons With Prior Work

The research protocol builds on existing knowledge in the literature. We are unaware of any published national Australian frameworks or models that directly relate to responding to women and pets fleeing domestic violence. Previous literature indicates when women are seeking help to flee from domestic violence, the risk of safety increases for both women and their pets. In addition, the pet may be used as a coercive control measure, risking further abuse for both the woman and the animal [18]. The evidence indicates it is vital to address the increased risks to safety when fleeing domestic violence. The prospective data collection of service providers and women using domestic violence and animal welfare services in Australia, as we propose in this study, enables further understanding and development of an Australian framework that is embedded by those with lived experiences to improve outcomes.

Strengths and Limitations

Limitations include the small sample sizes that will not be generalizable to the wider populations, and the exclusion of non-English-speaking populations limits the ability of the research to understand the special needs of the linguistic and cultural populations [55]. The primary target population is recruited from service providers and is considered safe to participate. This is a limitation for women and pets in situations that did not seek formal service provision, had stayed in the relationship, or were not safe from abuse. Bias is more likely to occur in qualitative research than in quantitative methods, resulting in difficulty reaching true objectivity [59]. However, the strength of the qualitative design allows for flexibility and
sensitivity in language, trust, rapport building, exploration of experiences, and collaboration within the community [58] and is appropriate for the study’s aims.

Conclusions
A research project guided by a conceptual framework informed by the findings of the scoping review confirms 4 key studies required to better understand the strengths, needs, and gaps of existing policy and support services for women and pets fleeing domestic violence, and the impacts of forced separation from companion animals. Ultimately, the project will develop an Australian national framework that will develop and provide more relevant guidance for supporting women and their pets fleeing domestic violence situations to improve outcomes for both women and their companion animals in Australia.

Acknowledgments
The authors would like to acknowledge Jacqui Watt for her expert advice on domestic violence; the advisory panel members for their contributions regarding animal welfare and domestic violence; and Anna Ludvik, Bridget Brown, Mandy Paterson, Marisa Debattista, and Susan Elvery and the Human Research Ethical Committee and the University Librarians for their contributions. Mentoring support was provided by Diana Mendez and the University Doctoral Cohort Program.

Data Availability
Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Authors’ Contributions
As the first author, JM developed the first draft of the manuscript guided by the senior author ZL. Both ZL and JL contributed to writing all sections of the manuscript and critically reviewed and approved the final version.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Description of the scoping review articles by separation event. [DOCX File, 33 KB - resprot_v13i1e52067_app1.docx ]

References


55. Pedemont K. Excluding non-English speaking people from health research including falls research for community-dwelling older people. Aust Orthopt J 2018;50:26-30 [FREE Full text]


Corrigenda and Addenda

Correction: Mobile Phone Technology for Preventing HIV and Related Youth Health Problems, Sexual Health, Mental Health, and Substance Use Problems in Southwest Uganda (Youth Health SMS)- Protocol for a Pilot Randomized Controlled Trial

Philip Kreniske¹, PhD; Olive Imelda Namuyaba², BA; Robert Kasumba³, MS; Phionah Namatovu² ³, MPH; Fred Ssewamala² ³, PhD; Gina Wingood⁴, SCD; Ying Wei⁵, PhD; Michele L Ybarra⁶, PhD; Charlotte Oloya⁴, LLB, LLM; Costella Tindyebwa⁷, MA; Christina Ntulo⁸, MA; Vincent Mujune⁷, MPH; Larry W Chang⁹, MPH, MD; Claude A Mellins¹⁰, PhD; John S Santelli¹¹, MPH, MD

¹Community Health and Social Sciences Department and The Institute for Implementation Science in Population Health, Graduate School of Public Health and Health Policy, City University of New York, New York, NY, United States
²International Center for Child Development, Masaka, Uganda
³Washington University in St Louis, St Louis, MO, United States
⁴Department of Sociomedical Sciences, Mailman School of Public Health, Columbia University, New York, NY, United States
⁵Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, United States
⁶Center for Innovative Public Health Research, San Clemente, CA, United States
⁷StrongMinds Uganda, Kampala, Uganda
⁸Malachite Center for Mental Health, Kampala, Uganda
⁹HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute and Columbia University, New York, NY, United States
¹⁰Heilbrunn Department of Population and Family Health, Mailman School of Public Health, Columbia University, New York, NY, United States

Corresponding Author:
Philip Kreniske, PhD
Community Health and Social Sciences Department and The Institute for Implementation Science in Population Health, Graduate School of Public Health and Health Policy, City University of New York
55 W 125th Street
New York, NY, 10027
United States
Phone: 1 6463649600
Email: philip.kreniske@sph.cuny.edu

Related Article:
Correction of: https://www.researchprotocols.org/2023/1/e49352
doi:10.2196/55725

In “Mobile Phone Technology for Preventing HIV and Related Youth Health Problems, Sexual Health, Mental Health, and Substance Use Problems in Southwest Uganda (Youth Health SMS): Protocol for a Pilot Randomized Controlled Trial” (JMIR Res Protoc 2023;12:e49352), the authors noted two errors:

The affiliation of author Claude A Mellins was:

Community Health and Social Sciences Department,
Graduate School of Public Health and Health Policy,
City University of New York, New York, NY, United States

It has been revised to:

HIV Center for Clinical and Behavioral Studies,
New York State Psychiatric Institute and Columbia University,
New York, NY, United States

The affiliation of authors Charlotte Oloya, Costella Tindyebwa and Vincent Mujune was:

Malachite Center for Mental Health, Kampala, Uganda

It has been revised to:

StrongMinds Uganda, Kampala, Uganda

The correction will appear in the online version of the paper on the JMIR Publications website on January 8, 2024, together with the publication of this correction notice. Because this was
made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 21.12.23; this is a non–peer-reviewed article; accepted 21.12.23; published 08.01.24.

Please cite as:
Correction: Mobile Phone Technology for Preventing HIV and Related Youth Health Problems, Sexual Health, Mental Health, and Substance Use Problems in Southwest Uganda (Youth Health SMS)- Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2024;13:e55725
URL: https://www.researchprotocols.org/2024/1/e55725
doi:10.2196/55725
PMID:
Corrigenda and Addenda

Correction: A Machine Learning Model to Predict Patients’ Adherence Behavior and a Decision Support System for Patients With Metastatic Breast Cancer: Protocol for a Randomized Controlled Trial

Marianna Masiero1,2, PhD; Gea Elena Spada2, MSc; Virginia Sanchini1, PhD; Elisabetta Munzone3, MD; Ricardo Pietrobon4, PhD; Lucas Teixeira4, MSc; Mirtha Valencia4, MSc; Aline Machiavelli1, MSc; Elisa Fragale2, MSc; Massimo Pezzolato1,2, MSc; Gabriella Pravettoni1,2, PhD

1Department of Oncology and Hemato-oncology, University of Milan, Milan, Italy
2Applied Research Division for Cognitive and Psychological Science, European Institute of Oncology IRCCS, Milan, Italy
3Division of Medical Senology, European Institute of Oncology IRCCS, Milan, Italy
4SporeData Inc, Durham, NC, United States

Corresponding Author:
Marianna Masiero, PhD
Department of Oncology and Hemato-oncology
University of Milan
Via Festa del Perdono 7
Milan, 20122
Italy
Phone: 39 0294372099
Email: marianna.masiero@unimi.it

Related Article:
Correction of: https://www.researchprotocols.org/2023/1/e48852
doi:10.2196/55928

In “A Machine Learning Model to Predict Patients’ Adherence Behavior and a Decision Support System for Patients With Metastatic Breast Cancer: Protocol for a Randomized Controlled Trial” (JMIR Res Protoc 12(1); e48852) the authors noted one error.

In the Acknowledgments, the last sentences appeared as follows:

The 8-item Morisky Medication Adherence Scale content, name, and trademarks are protected by US copyright and trademark laws. Permission for use of the scale and its coding is required. A license agreement is available from MMAS, LLC. The property/identifier issue/registration date/categories are as follows: US registration 5,837,273 August 20, 2019, Trade/Service Mark; US registration TX-8-285-390 June 12, 2016, Copyright; and US registration TX-8-632-533 September 21, 2018, copyright.

This project has been funded by a Pfizer grant: Enhancing therapy adherence among patients with metastatic breast cancer (65080791).

This has been changed to appear as follows:

The MMAS-8 Scale, content, name, and trademarks are protected by US copyright and trademark laws. Permission for use of the scale and its coding is required. A license agreement is available from MMAR, LLC., www.moriskyscale.com.

This project has been funded by a Pfizer grant: Enhancing therapy adherence among patients with metastatic breast cancer (65080791).

The correction will appear in the online version of the paper on the JMIR Publications website on January 15, 2024 together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.
Correction: A Machine Learning Model to Predict Patients’ Adherence Behavior and a Decision Support System for Patients With Metastatic Breast Cancer: Protocol for a Randomized Controlled Trial

JMIR Res Protoc 2024;13:e55928
URL: https://www.researchprotocols.org/2024/1/e55928
doi:10.2196/55928
PMID:38224582
Correction: The Development of a UK Culturally Adapted and Modified Version of the Person Attuned Musical Interactions Manual: Protocol for a 2-Phase Mixed Methods Study

Bryony Waters¹, BSc, MSc; Martin Orrell¹, BMedSci, MBBS, PhD; Orii McDermott¹, ARCM, BA, MMT, PhD

Institute of Mental Health, School of Medicine, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Bryony Waters, BSc, MSc
Institute of Mental Health
School of Medicine
University of Nottingham
Jubilee Campus
Innovation Park, Triumph Road
Nottingham, NG7 2TU
United Kingdom
Phone: 44 1158231294
Email: bryony.waters@nottingham.ac.uk

Related Article:
Correction of: https://www.researchprotocols.org/2023/1/e43408

(JMIR Res Protoc 2024;13:e56085) doi:10.2196/56085


1. The title “The Development of a UK Culturally Adapted Version of the Person Attuned Musical Interactions Manual: Protocol for a 2-Phase Mixed Methods Study” has been revised to:

   The Development of a UK Culturally Adapted and Modified Version of the Person Attuned Musical Interactions Manual: Protocol for a 2-Phase Mixed Methods Study

2. PAMI-United Kingdom was revised to PAMI-Modified, with all subsequent mentions in the paper being changed from PAMI-UK to PAMI-M.

3. In Table 1, all instances of the word Dutch were revised to Danish, as the language of the intervention is Danish.

4. Also in Table 1, the phrases in the title and footnote PAMI-DK and PAMI-DK- Person Attuned Musical Interaction Danish Version have been replaced by PAMI and PAMI-Person Attuned Musical Interactions, respectively.

5. In the Corresponding Author’s address, Triumph Road was corrected to Triumph Road.

6. Some sentences were revised to add the words modified, modification or modifying where necessary, as follows:

   In the Abstract, in the section Background, “…a team of researchers in the United Kingdom have culturally adapted the tool.” has been edited to:

   …a team of researchers in the United Kingdom have modified and culturally adapted the tool.

   In the Abstract, in the section Objective, “This study aims to investigate the appropriateness of the adapted UK manual for UK care homes and to explore…” has been edited to:

   This study aims to investigate the appropriateness of the adapted and modified manual for UK care homes and to explore...

   In the Abstract, in the section Conclusions, “This study will be the first to investigate the culturally adapted UK PAMI” has been edited to:

   This study will be the first to investigate the modified version of PAMI.

   In Background, in the section Person Attuned Musical Interactions Manual, “a UK version of PAMI was developed by a team of researchers at the University of Nottingham” has been edited to:

   A modified version of PAMI was developed by a team of researchers at the University of Nottingham.

   In Background, in the section Person Attuned Musical Interactions Manual, “The first author’s thesis reports the manual development process and the differences between the Danish and UK version of PAMI” has been edited to:
The first author’s thesis reports the manual development process and the differences between the Danish and Modified version of PAMI.

In Methods, under the section Training Session, “BW was the researcher who developed the translated PAMI manual as part of her PhD” has been edited to:

BW was the researcher who developed the modified PAMI manual as part of her PhD.

In Methods, under the section Intervention Development, “...the research team liaised with the Danish PAMI team to ensure that any adaptation remained consistent with the PAMI ethos” has been edited to:

...the research team liaised with the Danish PAMI team to ensure that any adaptation and modification remained consistent with the PAMI ethos.

In Discussion, under the section Anticipated Findings, “This study will be the first to investigate the culturally adapted PAMI. Therefore, the data collected in phases 1 and 2 will provide feedback on the appropriateness of the manual for staff...” has been edited to:

This study will be the first to investigate the modified PAMI-M. Therefore, the data collected in phases 1 and 2 will provide feedback on the appropriateness of the modified manual for care staff...

In Discussion, under the section Anticipated Findings, “This study will investigate whether the UK-adapted version of PAMI can benefit residents,....” has been edited to:

This study will investigate whether the modified version of PAMI can benefit residents,...

In Conclusions, the phrase “The study will be the first to investigate the culturally adapted UK PAMI” has been edited to:

The study will be the first to investigate the modified PAMI.

In Conclusions, “...and explore participants’ experience of using PAMI” has been edited to:

...and explore participants’ experience of using a modified version of PAMI.

In the Caption of Figure 2, “PAMI: Person Attuned Musical Interactions” has been edited to:

PAMI-M: Person Attuned Musical Interactions-Modified.

In the caption of Figure 3, the phrase “The manual development process for culturally translating and adapting the Danish Person Attuned Musical Interactions (PAMI) intervention for UK care homes. PAMI-DK: Person Attuned Musical Interactions Danish version; PAMI-UK: Person Attuned Musical Interactions United Kingdom” has been edited to:

The manual development process for culturally translating and modifying the Person Attuned Musical Interactions (PAMI) intervention for UK care homes. PAMI: Person Attuned Musical Interactions.

In the caption of Table 1, “The changes between the Danish and UK Person Attuned Musical Interactions (PAMI)” has been edited to:

The changes between the Danish and Modified Person Attuned Musical Interactions (PAMI).

The correction will appear in the online version of the paper on the JMIR Publications website on January 26, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 05.01.24; this is a non–peer-reviewed article; accepted 05.01.24; published 26.01.24.

Please cite as:
Waters B, Orrell M, McDermott O
Correction: The Development of a UK Culturally Adapted and Modified Version of the Person Attuned Musical Interactions Manual: Protocol for a 2-Phase Mixed Methods Study
JMIR Res Protoc 2024;13:e56085
URL: https://www.researchprotocols.org/2024/1/e56085
doi:10.2196/56085
PMID:
Abstract

Background: The last 2 decades have been a time of exponential growth and maturation for digital health, while the global burden of respiratory disease continues to grow worldwide. Leveraging digital health interventions (DHIs) to manage and mitigate respiratory disease and its adverse health effects presents itself as an obvious path forward.

Objective: We aimed to understand the current digital landscape and enabling environment around respiratory health to reduce costs, avoid duplication, and understand the comprehensiveness of DHIs.

Methods: This study will follow a scoping review methodology as outlined by Arksey and O’Malley, the Joanna Briggs Institute, and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist. MEDLINE, Embase, CINAHL, PsycINFO, Cochrane Library, Web of Science, PakiMedNet, and MyMedR databases will be searched along with key websites, repositories, and gray literature databases. The terms “respiratory health,” “digital health,” “South Asia,” and “Southeast Asia,” as well as related terms will be searched. The results will be screened for duplicates and then against the inclusion and exclusion criteria. For the studies included, data will be extracted, collated, and analyzed.

Results: The scoping review was started in July 2023 and will be finalized by February 2024. Results will be presented following the World Health Organization’s classification of DHIs to categorize interventions in a standardized format and the mobile health evidence reporting and assessment checklist to report on the effectiveness of interventions. Further exposition of the evidence extracted will be presented through narrative synthesis.

Conclusions: As DHIs continue to proliferate, the need to understand the current landscape becomes more pertinent. In this scoping review, we will seek to more clearly understand what digital health tools and technologies are being used in the current landscape of digital health in South and Southeast Asia for respiratory health and to what extent they are addressing the respiratory health needs of the region. The results will inform recommendations on digital health tools for respiratory health in South and Southeast Asia will help funders and implementers of DHIs leverage existing technologies and accelerate innovations that address documented gaps in the studied countries.

International Registered Report Identifier (IRRID): DERR1-10.2196/52517

(JMIR Res Protoc 2024;13:e52517) doi:10.2196/52517

KEYWORDS
digital health; respiratory health; Asia; scoping review; landscape mapping; digital health intervention; digital health environment; mobile health; mHealth
**Introduction**

Digital health and care refer to the use of information communication technologies, commonly known as digital health interventions (DHIs), by health and care professionals, patients, carers, or health managers to manage illnesses and wellness [1-4]. The last 2 decades have been a time of exponential growth and maturation for digital health [4] due to the promises of improved health for all and personal health empowerment [5]. Concurrently, the global burden of respiratory disease continues to grow worldwide, with infectious and noncommunicable respiratory diseases being among the top 10 medical conditions (out of 369 diseases and injuries measured) in terms of years of life lost due to premature death and years lived with a disability (measured by disability-adjusted life years) [6,7]. This increased burden of respiratory disease is more acutely felt in Asia, where mortality rates are higher and public awareness and government engagement are lower than in other regions of the world [8,9]. Leveraging DHIs to manage and mitigate respiratory disease and its adverse health effects presents itself as an obvious path forward. However, the first step to harnessing the power of digital health must be understanding the current digital landscape and enabling the environment to reduce costs, avoid duplication, and increase the efficiency, accessibility, and sustainability of such interventions [10-13].

The National Institute for Health Research (NIHR)–funded Global Health Research Unit on Respiratory Health (RESPIRE) is a collaboration between several Asian organizations and universities in Bangladesh, Bhutan, India, Malaysia, Pakistan, Indonesia, and Sri Lanka and the University of Edinburgh in Scotland, United Kingdom [9,14,15]. RESPIRE aims to achieve the following:

- develop into a world-leading Unit that will: (i) map and collate continuing and emerging respiratory challenges; (ii) prioritise existing evidence-based interventions that have the potential to be adapted to reduce mortality and morbidity in low- and middle-income countries (LMICs); (iii) support local adaptation and tailoring of interventions for deployment in low-resource environments and catalyse developmental work in areas of unmet need; (iv) support local implementation efforts and evaluation of programmes of work; and (v) identify the best delivery mechanisms for long-term delivery and scaling-up.

This is done through 4 different translational platforms, of which 1 platform focuses on Digital Health and Innovation and aims to provide holistic support to partner countries on the design, funding, deployment, and sustainability of new and existing digital health interventions for respiratory health. This scoping review will contribute to advancing the aims and work of Digital Health and Innovations.

Understanding the current landscape of DHIs that target respiratory diseases will (1) uncover existing gaps, (2) highlight potential opportunities, (3) suggest research and program priorities most needed in the field of digital health to address current respiratory health diseases in South and Southeast Asia, and (4) further advance RESPIRE’s overall aims. Therefore, we aim to undertake a scoping review to map respiratory disease DHIs in South and Southeast Asia to identify existing technologies, opportunities, and gaps, and to put forward pertinent recommendations from the insights gained.

**Methods**

**Scoping Review Methodology**

The scoping review methodology, as outlined by Arksey and O’Malley and the Joanna Briggs Institute (JBI) [16,17], is an appropriate approach for mapping DHIs. Scoping reviews allow flexibility when exploring the diverse and heterogenous field of digital health, are appropriate when using different sources of data (eg, peer-reviewed journals, gray literature, and expert opinions), and permit inclusion and exclusion criteria to be iteratively refined as more evidence is uncovered [16-19]. Additionally, the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [20] will be followed to ensure adherence to methodological standards.

**Identifying the Research Question**

This scoping review will seek to answer the following research questions:

1. What digital health tools and technologies are being used in South and Southeast Asia for respiratory health?
2. How are these addressing (or not) the respiratory health needs of the region?
3. What recommendations can be made from the literature?

**Identifying Relevant Studies**

To identify relevant literature, academic databases will be searched along with key websites, repositories, and gray literature databases that may contain relevant information for our scoping review. Textbox 1 shows the proposed databases.

The search strategy has been initially drafted for MEDLINE in the Ovid platform (Textbox 2) and will be adapted for the remaining databases. The search strategy has been iteratively developed and refined by the authors’ input and the librarian at the University of Edinburgh. The terms “respiratory health,” “digital health,” “South Asia,” and “Southeast Asia,” as well as all relevant variations of these terms have been included in the search strategy to gather as much pertinent literature and information as possible.
Textbox 1. Proposed database and key websites.

**Databases**
- MEDLINE
- Embase
- CINAHL
- PsycINFO
- Cochrane Library
- Web of Science
- PakMediNet
- MyMedR

**Other sources**
- ProQuest Thesis and Dissertations
- Digital Health Atlas
- Global Digital Health Index
- Digital Square’s Map and Match
- World Health Organization’s Global Index Medicus

Textbox 2. Search strategy.

**Search terms**

1. occupational diseases/ or Asthma/ or Air Pollution/ or Respiratory Tract Diseases/ or Occupational Exposure/ or Air Pollutants/ or Tuberculosis/ or Tuberculosis, Pulmonary/ or "Tobacco Use"/ or Tobacco/ or "Tobacco Use Cessation"/ or "Tobacco Use Cessation Devices"/ or Pulmonary Disease, Chronic Obstructive/ or Pneumonia/ or COVID-19/

2. (respiratory health or tuberculosis or tobacco or pneumonia).mp

3. 1 or 2

4. telemedicine/ or telehealth/ or artificial intelligence/ or machine learning/ or medical informatics/ or electronic health records/ or mobile applications/ or exp Informatics/

5. (artificial intelligence or digital health or e-health or ehealth or mhealth or m-health or (digital adj2 (health or solution* or system*)) or (health adj2 (electronic or record* or tele*)) or ict4d or (information adj5 development) or machine learning or mobile health or telecare or telehealth or telemedicine or tele-health or teleconsultation or tele-consultation or tele-care or tele-medicine or (tele adj1 (medicine or care or health or consultation)) or ((virtual* or remote*) adj4 (visit* or consult* or meet* or appoint* or communicat*)) or (Health* adj4 tech*) or e-portal* or eportal* or (Patient* adj2 portal*) or (medical adj2 informatic*).mp.

6. 4 or 5

7. Asia/ or Asia, Southern/ or Asia, Southeastern/

8. (Asia or Brunei Darussalam or Cambodia or Indonesia or Lao People’s Democratic Republic or Malaysia or Myanmar or Philippines or Singapore or Thailand or Timor-Leste or Viet Nam or Vietnam or Afghanistan or Bangladesh or Bhutan or India or Iran or Islamic Republic of Iran or Maldives or Nepal or Pakistan or Sri Lanka).mp

9. 7 or 8

10. 3, 6, and 9

**Study Selection**

Exclusion and inclusion criteria have been developed according to the domains of population, concept, context, and type of evidence, suggested by the JBI [17,18]. Additionally, an “other variables” category has been created to include year, language, and format criteria (Table 1). The regions of South and Southeast Asia include 19 countries as established by the United Nations Statistics Division [21]. These 2 regions have been chosen because they contain all the RESPIRE2 countries. Multicountry studies containing countries from the selected regions and other regions of the world will be included in the initial screening and only excluded if they do not provide the data of interest separate for each country. Only studies in English and those published in the last 10 years (since 2013) will be included to keep the scope of this review within manageable boundaries.
Table 1. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Any population</td>
<td>N/A&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Concept</td>
<td>Technological interventions for respiratory health that fall under any of the categories of the World Health Organization’s classification of digital health interventions</td>
<td>Other nontechnological interventions used for respiratory health</td>
</tr>
<tr>
<td></td>
<td>Respiratory health includes respiratory infections, non-communicable respiratory diseases, and preventable risk factors for respiratory conditions, as defined by RESPIRE&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Not a respiratory health focus</td>
</tr>
<tr>
<td>Context</td>
<td>South and Southeast Asia</td>
<td>The rest of the world</td>
</tr>
<tr>
<td>Types of evidence sources</td>
<td>Any study design</td>
<td>Books, Abstracts only, Posters, Protocols</td>
</tr>
<tr>
<td>Other variables</td>
<td>Published in English</td>
<td>Published in any other language</td>
</tr>
<tr>
<td></td>
<td>Studies or data published in the last 10 years (2013-2023)</td>
<td>Studies or data published before 2013</td>
</tr>
<tr>
<td></td>
<td>Full article or data are available digitally</td>
<td>Full article or data are not available digitally</td>
</tr>
</tbody>
</table>

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

<sup>b</sup>RESPIRE: The National Institute for Health Research–funded Global Health Research Unit on Respiratory Health.

For information found in a scientific study format, Covidence software [22] will be used to eliminate duplicates and carry out screening and extraction. After deduplication, title, abstract, and full-text screening will be done by 2 authors according to the inclusion and exclusion criteria. Discrepancies will be first addressed by consensus between those 2 authors. If there is a lack of consensus, a third reviewer will decide.

For all other types of information or data found from searches, manual screening by 1 reviewer will happen first. Relevant data will be entered into a spreadsheet, and a second reviewer will assess it. Discrepancies will first be addressed between both reviewers, and if there is no consensus, a third reviewer will make the final decision. We will not contact authors directly to further understand whether a study should be included since it would most likely significantly lengthen the timeline for this scoping review.

Scoping reviews use secondary data and do not require ethics approval under RESPIRE rules. However, the authors will adhere to the highest ethical standards when carrying out the review. This protocol establishes a transparent and reproducible study design, which limits the potential for personal bias [23].

**Charting the Data**

After selection, relevant data will be extracted to a spreadsheet using Covidence. The extraction form will be first piloted in 3-5 studies to assess if it is fit for purpose as recommended by the JBI. Data selected from a nonscientific study format will be entered into the extraction spreadsheet as accurately as possible. However, it is acknowledged that not all fields may contain relevant information, and some fields may need to be modified (eg, data from the Digital Health Atlas will have a “source of information” field instead of an “author” field).

When 2 or more articles refer to the same overall study, those articles will be grouped as one before data extraction.

**Collating, Summarizing, and Reporting the Results**

After data analysis, the data will be collated and analyzed as follows:

- Quantitative data (ie, the number of studies, type of study, and year) will be presented.
- Narrative synthesis will be used to answer the research questions and to present further data extracted.

**Consultation**

Consultation with stakeholders will be ongoing throughout the scoping review process. We will disseminate early findings among partners so that they can provide feedback on findings and that feedback can be incorporated into the discussion.

**Results**

This scoping review was started in July 2023 and will be finalized by February 2024. Preliminary findings will be shared with stakeholders and a final write-up of the scoping review is projected to be finalized by the end of March 2024. To date, 10,469 studies have been screened. The screening of abstracts is underway.
**Discussion**

**Expected Outcomes**

Through this scoping review, we will seek to better understand what digital health tools and technologies are being used in South and Southeast Asia for respiratory health and to what extent they are addressing the respiratory health needs of the region. The results will inform recommendations on digital health tools for respiratory health in South and Southeast Asia and will help funders and implementers of DHIs leverage existing technologies and accelerate innovations that address documented gaps in the studied countries. The results of this review will be limited by the fact that only studies in English and studies published in the last 10 years will be included. This review will enhance the knowledge of digital health tools and technologies in the region, which is paramount before undertaking any new initiative, as it helps prevent redundant work and investment by building on existing systems and lessons learned.

**Conclusions**

As DHIs, in general and in respiratory health in particular, continue to proliferate, the need to understand the current landscape becomes more pertinent. Through this scoping review, we will systematically map out DHIs, which serves as the required first step in any well-informed and thought-out design and deployment of DHIs.

**Acknowledgments**

We would like to thank Marshall Dozier, the librarian at the University of Edinburgh, who provided feedback on the protocol’s search strategy and overall design.

Generative artificial intelligence was not used at any time during the research and writing of this protocol.

This research was funded by the UK National Institute for Health and Care Research (NIHR)—Global Health Research Unit on Respiratory Health (RESPIRE; NIHR132826)—using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the United Kingdom Government.

**Data Availability**

The data resulting from this scoping review will be made available as supplementary materials at the time of publication.

**Authors’ Contributions**

JE and LE conceptualized the idea. LE drafted the manuscript and search strategy. JE, MF, AA, and ZA provided feedback on the draft.

**Conflicts of Interest**

AA is a shareholder of a digital health company in Malaysia (UMCH Technology Sdn Bhd).

**References**

Abbreviations

DHI: digital health intervention
JBI: Joanna Briggs Institute
NIHR: National Institute for Health Research
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RESPIRE: National Institute for Health Research–funded Global Health Research Unit on Respiratory Health


14. NIHR RESPIRE. IPCRG. URL: https://www.ipcrg.org/projects/research/nihr-respire [accessed 2023-03-16]

15. What is RESPIRE? University of Edinburgh. URL: https://www.ed.ac.uk/usher/respire/about/what-is-respire [accessed 2023-03-16]


Protocol


Mark Groulx1, PhD; Shannon Freeman2, PhD; Keone Gourlay3, BPI; Dawn Hemingway4, PhD; Emma Rossnagel1, MPH; Habib Chaudhury5, PhD; Mohammadjavad Nouri5, PhD

1University of Northern British Columbia, Prince George, BC, Canada
2School of Nursing, University of Northern British Columbia, Prince George, BC, Canada
3School of Planning and Sustainability, University of Northern British Columbia, Prince George, BC, Canada
4School of Social Work, University of Northern British Columbia, Prince George, BC, Canada
5Department of Gerontology, Simon Fraser University, Vancouver, BC, Canada

Corresponding Author:
Mark Groulx, PhD
University of Northern British Columbia
3333 University Way
Prince George, BC, V2N 4Z9
Canada
Phone: 1 250 960 5837
Email: mark.groulx@unbc.ca

Abstract

Background: The number of people in society living with dementia is growing. In Canada, most people who live with dementia live at home, often in a neighborhood setting. Neighborhood environments can be a source of independence, social engagement, and well-being. They can also contain barriers that limit physical activity, social engagement, and well-being. A dementia-friendly neighborhood includes assets that support persons living with dementia and their caregivers in multiple life domains, including those that support walking within the neighborhood environment.

Objective: The objectives for this scoping review are twofold. First, focusing on walkshed analysis, we aim to extend scholarly understandings of methodological practices used in the monitoring and evaluation of dementia-friendly neighborhoods. Second, we aim to provide clear and practical guidance for those working in planning, design, and public health fields to assess the neighborhood context in support of evidence-based action to improve the lives of persons living with dementia.

Methods: The study design follows Arksey and O’Malley’s scoping review framework and PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) guidelines. We will conduct a search of peer-reviewed studies in 6 electronic databases to identify the use of Geographic Information System analysis to measure the walkshed of persons living with dementia in a community setting. As age is a primary risk factor associated with dementia, we will also include studies that focus more broadly on community-dwelling older adults aged 65 years and older. Data will be extracted, analyzed, and represented according to 3 domains. This includes study details, walkshed analysis methods, and criteria and indicators used to measure dementia-friendly neighborhoods.

Results: The results of the study and the submission of a manuscript for peer review are expected in June 2024. The results of the review are expected to contribute to an understanding of methods for monitoring and evaluating dementia-friendly neighborhoods. Expected findings will include a detailed breakdown of current parameters and routines used to conduct walkshed analysis. Findings will also convey criteria that can be operationalized in a Geographic Information System as indicators to assess barriers and facilitators to walking in a neighborhood setting.

Conclusions: As far as we are aware, the proposed scoping review will be the first to provide comprehensive methodological or technical guidance for conducting walkshed analysis specific to persons living with dementia. Both the scalability and objective nature of walkshed analysis are likely to be of direct interest to public health practitioners, planners, and allied professionals. Clearly documenting methods used in walkshed analysis can spur increased collaboration across these disciplines to enable an evidence-informed approach to improving neighborhood environments for persons living with dementia.

International Registered Report Identifier (IRRID): PRR1-10.2196/50548
**Introduction**

**Overview**

Walkshed analysis identifies the extent of the community environment surrounding a central location that is accessible at a scale where walking is a competitive mode of mobility [1]. Once a walkshed is delineated in a Geographic Information System (GIS), criteria and indicators can identify barriers and enablers to walking [2]. Walkshed analysis is relevant to planning and public-health partnerships that seek to support persons living with dementia. More than 55 million people are currently living with dementia across the world. The global prevalence of dementia is projected to continue to rise by approximately 10 million new cases per year [3]. In Canada, most people experiencing dementia live at home. As of 2016, around 69% of those aged 80 years or younger were living outside of the long-term care system [4]. As an umbrella term, dementia captures the experience of progressive cognitive decline. It can impact an individual’s mood, behavior, and actions, including the performance of key activities of daily living [5]. There are many types of dementia, including Alzheimer dementia, vascular dementia, frontotemporal dementia, Lewy body dementia, mixed dementia, and young-onset dementia. Alzheimer disease is the most common cause, contributing to 60% to 70% of cases of dementia [6].

Literature on dementia-friendly communities (and neighborhoods) takes a relational view [7,8]. This view acknowledges that well-being is conditioned by interrelated aspects of a person’s social, built, and ecological surroundings [8-10]. Accordingly, scholars identify dementia-friendly environments as the arrangement of supportive assets into a community fabric that promotes meaningful societal engagement for persons living with dementia and their caregivers [11]. This includes the complex social relations that persons living with dementia experience in a community setting, making the physical neighborhood part of a relational and moral context [12,13].

Scholarship on dementia-friendly communities and neighborhoods stems from calls to better support persons who are living with dementia outside of an institutional setting [11]. These calls reflect the fact that scholars have long viewed neighborhoods as a central relational context shaping individual behavior and life quality [14]. As early as the turn of the 20th century, ideas about neighborhood planning in North America drew on sociological concepts such as Charles H Cooley’s primary group. The primary group and similar concepts asserted that the neighborhood was the main setting for the social relations that informed one’s perspectives and ideals [15].

A long fascination with neighborhood environments helps explain the growing effort to understand how the neighborhood can enable or hinder self-determination for persons living with dementia. This includes aspects of identity development and one’s ability to shape life balance [10,12,16]. Remaining close to the home, or aging in place, is also “closely intertwined with (a person’s) sense of self and identity” [17]. By contrast, moving away from familiar areas can have negative effects on persons living with dementia [18]. To remain active and engaged within their environments while aging in place, persons living with dementia need special considerations and support in their neighborhoods [17].

The influence of the built environment on a sense of community and one’s place therein remains up for debate in an increasingly mobile and digital society [19,20]. At the same time, there is a convincing body of evidence demonstrating that planning and design can impact behavior. The extent to which a neighborhood setting encourages or discourages important social and health behaviors such as walking is a particular focus for planning-health partnerships [21-23]. There is also a growing body of evidence illustrating that walking outdoors boosts quality of life for those living with dementia, contributing to improved mood, quality of sleep, and sense of freedom [17,24,25].

Urban planning scholar Lawrence Frank significantly advanced the conception and measurement of walkability. He describes walkability as the extent to which an environment’s social and physical characteristics promote walking as a competitive and desirable form of mobility [26,27]. Recent work has extended the idea of walkability to a more encompassing notion of “active living environments.” Active living environments are defined as “the emergent natural, built, and social properties of neighborhoods that promote physical activity and health and allow for equitable access to health-enhancing resources” [28].

Scholars have used a wide variety of methods to study walkability and its relation to walking behavior. These include phenomenological interviews [29], cross-sectional community surveys [30], observational techniques [31], surveys [32,33], photovoice [34,35], and in-situ walking interviews [12]. Scholars have also deployed criteria and indicators that enable monitoring and evaluation of the social, built, and ecological environments that make up a city [36]. In some cases, criteria and indicators are operationalized using a geospatial approach that assesses barriers and facilitators to walking in a small area (eg, 1 km) surrounding a central location such as a residence. This approach is often referred to as walkshed analysis.

In North America, walkability is now well researched within urban settings in the context of the “general population.” By comparison, factors that shape walkability for members of equity-deserving groups, particularly persons living with dementia, are comparatively understudied. There is a need to better document (1) what walkability criteria and indicators are relevant to the lived experience of persons living with dementia, (2) how methods are operationalized to examine barriers and facilitators using a walkshed approach, and (3) where barriers and facilitators of walkability for persons living with dementia...
may align or conflict with those of other populations. Given these needs, the objectives for this scoping review are twofold:

1. Focusing on walkshed analysis, extend scholarly understandings of methodological practices used in the monitoring and evaluation of dementia-friendly neighborhoods.

2. Provide clear and practical guidance for those working in planning, design, and public health fields to assess the neighborhood context in support of evidence-based action to improve the lives of persons living with dementia.

To achieve the preceding objectives, this scoping review will address the following research question: What dimensions, criteria, and indicators can be recognized within the academic literature for measuring neighborhood walkability for persons living with dementia based on a walkshed methodology?

### Existing Reviews

This protocol was informed by an initial review of existing peer-reviewed literature. The purpose of this review was to identify possible knowledge syntheses on the use of walkshed methodology to document barriers and facilitators faced by persons living with dementia. Table 1 summarizes key aspects of 6 related knowledge syntheses. All but 1 of the identified studies were published within the past 5 years [36]. A total of 2 of the studies directly focused on persons living with dementia. Other studies focused on dementia risk factors among older adults (see Table 1).

#### Table 1. Summary of comparable existing knowledge syntheses as they relate to the proposed scoping review.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
<th>Objective</th>
<th>Population focus</th>
<th>Addresses aspects of walkshed methods</th>
<th>Addresses objective criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akinci et al</td>
<td>How different are objective operationalizations of walkability for older adults compared to the general population? a systematic review</td>
<td>Summarize and compare methods used to operationalize objective walkability for older adults and the general population</td>
<td>Older adults or general population</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cerin et al</td>
<td>The neighbourhood physical environment and active travel in older adults: a systematic review and meta-analysis</td>
<td>Identify correlates of neighborhood physical features and active travel in older adults and quantify the strength of associations</td>
<td>Older adults</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Sturge et al</td>
<td>Features of the social and built environment that contribute to the well-being of persons with dementia who live at home: a scoping review</td>
<td>Summarize evidence from qualitative studies about how social and built environment features influence well-being for persons living with dementia</td>
<td>Persons living with dementia</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Gan et al</td>
<td>Dementia-friendly neighbourhood and the built environment: a scoping review</td>
<td>Synthesize knowledge and support policy direction related to dementia-friendly neighborhood environments and attendant psychosocial outcomes</td>
<td>Persons living with dementia</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Peters et al</td>
<td>Measuring the association of objective and perceived neighborhood environment with physical activity in older adults: challenges and implications from a systematic review</td>
<td>Assess the correlates of neighborhood characteristics and physical activity in older adults to provide a body of evidence to support neighborhood environmental interventions</td>
<td>Older adults</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Chen et al</td>
<td>Neighbourhood-built environment associated with cognition and dementia risk among older adults: a systematic literature review</td>
<td>Assess the state of current knowledge on the links between neighborhood environments and cognitive health in older adults</td>
<td>Older adults at risk of dementia</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Gan and colleagues [25] reviewed 29 studies and documented methodologies ranging from applications of virtual reality to measurements of statistical association. No use of walkshed methods was reported. The authors also assessed the psychosocial outcomes of outdoor use (eg, increased social agency, anxiety, and promotion of personhood) and built environment characteristics that facilitate use and participation (eg, land use diversity, presence of landmarks, and irregular street grids).

By contrast, Sturge and colleagues [37] focused solely on qualitative studies exploring how social and built environments contribute to the well-being of persons living with dementia at home. Under a theme examining “connection to society and supportive relationships,” the authors review 4 key areas of support. These include contact with friends and family, social networks afforded by formal events and professional services, connections available across a host of neighborhood settings (eg, pubs and cafés), and the mixed reactions persons living with dementia can experience when disclosing their diagnosis. A second theme titled “interaction with natural environments and public space” examines supports (eg, parks and sounds of children playing) and barriers (eg, complex street environments and noise from traffic).

Both Peters and colleagues [2] and Akinci and colleagues [21] review (respectively) aspects of walkshed methodology in the context of older adults or older adults and the general public. Neither focused specifically on persons living with dementia. Peters and colleagues [2] distinguish between subjective and...
objective measures and discuss the use of accelerometers, GIS, and field-based audit approaches. They document key aspects related to the use of walkshed methods with older adults. Elements include operational definitions of a neighborhood, walking times or distances used to define a walkshed, and neighborhood attributes associated with walking and other physical activity. Akinci and colleagues [21] similarly report on GIS-based methods for spatial analysis. They report on walkshed buffer types and sizes and 167 different walkability variables across 24 studies of older adults.

The identified 6 studies are each related to the aim of this proposed scoping review. None directly cover the realm we seek to document. In 4 cases, the studies do not review objective walkshed methods. The remaining 2 cases do not focus on persons living with dementia.

**Methods**

**Study Eligibility**

The primary objective of this study is to report on research relevant to the use of walkshed methodology. We are specifically interested in walkshed analysis which involves the monitoring and evaluation of barriers and facilitators to walking in a neighborhood setting. Eligible studies will include those that reveal details about how to define a walkshed in a manner that is appropriate to the walking experience of persons living with dementia (eg, walking distance used to define a walkshed).

**Textbox 1.** Summary of the inclusion and exclusion process and the criteria (framed as prompts) used to exclude studies.

<table>
<thead>
<tr>
<th>Review level</th>
<th>Criteria</th>
</tr>
</thead>
</table>
| **Level 1:** title, abstract, and keyword review | - Does the study include a focus on geographic areas within a community setting?  
- Does the study include a focus on outdoor spaces?  
- Does the study include a focus on people’s use of the community environment by walking or other forms of non-motorized mobility? |
| **Level 2:** full text article review | - Did the study collect and analyze primary or secondary data following a structured methodological approach?  
- Does the study identify measurable criteria and indicators related to walkability or report on the use of walkshed methods?  
- Does the study specifically focus on environmental use by persons living with dementia or older adults? |

**Population and Setting**

This review will be guided by Arskey and O’Malley’s [39] 6-step scoping review process. It will include studies that involve participants recognized to be living with dementia or mild cognitive impairment and who reside in a community setting. Studies that focus on persons living in congregate care-based facilities such as assisted living homes and long-term care homes will be excluded. We expect to find few published studies that explicitly focus on this population in the context of operational aspects of walkshed methodology. As age is the primary risk factor associated with dementia, we will also include studies that focus more broadly on community-dwelling older adults aged 65 years and older [6]. We will track differences in existing evidence between these population groups.

**Search Strategy**

Our search strategy was developed by a project manager with experience conducting scoping reviews. It involved consultation with a research librarian and the broader research team. The latter consultation involved a workshop that iteratively identified, tested, and respaced search domains and terms. Our search strategy includes a combination of subject headings and title or abstract-focused keyword searching (Textbox 2). These strategies target the intersection of an activity or policy domain (walking), an environmental setting domain (outdoor neighborhood setting), and a population focus domain (persons living with dementia and older adults). We will apply search
strings to 6 electronic databases known to publish high-quality research around our focus domains (PubMed, Medline, CINAHL, APA PsycINFO, Business Source, and Web of Science). Endnote will be used to manage citations, and DistillerSR (DistilerSR Inc) and Excel (Microsoft Corporation) will be used to manage the inclusion, data extraction, and charting stages of this review.

**Textbox 2. Domain areas and search terms to be used in search strings for database searches.**

<table>
<thead>
<tr>
<th>Domain areas and search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Activity or policy - walking, walkshed, walkability, walk, wayfinding, way finding, indicator, criteria, dimension, requirement, experience, audit, measure</td>
</tr>
<tr>
<td>• Environmental setting - footpath, greenspace, green space, population density, rural population, neighbourhood characteristics, city planning, community*, neighbo<em>rhood</em>, built environment, urban design*, urban planning, town planning, city planning, building densit*, social densit*, population densit*</td>
</tr>
<tr>
<td>• Population focus - dementia, alzheimer*, aged</td>
</tr>
</tbody>
</table>

**Article Selection Process**

After removing duplicate sources from our initial study pool using DistillerSR, we will use DistillerSR to complete screening at 2 levels. At level 1, we will assess the title, abstract, and keywords of each potential source. This assessment will include 2 independent reviewers using the level 1 inclusion criteria in Textbox 1. Studies will be excluded if both reviewers definitively identify relevant content and answer no to any of the criteria prompts. Studies will be moved to level 2 screening if a prompt cannot be answered definitively. To promote consistency at level 2 article screening, 2 reviewers will assess the full text of all remaining sources. Studies will only be included if reviewers can definitively answer yes to all inclusion prompts. We will address discrepancies at each level at a team meeting that involves a reassessment of the source and a consensus decision made by the team.

**Data Charting and Representation**

Data charting and representation will follow 2 interrelated steps outlined by Arksey and O’Malley [39]. Common practices in scoping review methodology and existing knowledge syntheses documented above informed the creation of the data charting schema listed below. Using this schema, we will develop a data matrix in Excel. This matrix will organize data and allow for the analysis of key items of information. Following guidance from Levac and colleagues [40], we will review and iteratively update the initial schema shown in Table 2 as the final study pool is examined. A total of 2 reviewers will extract data for a subset of papers (n=5). They will compare and update the schema as they reflect on processes and outcomes. Final data extraction will be completed by a single reviewer.

**Table 2. Initial data charting schema for creation of data charting matrix.**

<table>
<thead>
<tr>
<th>Study details</th>
<th>Walkshed methods</th>
<th>Criteria and indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Definition of walkability</td>
<td>Measurement domains reported</td>
</tr>
<tr>
<td>Lead author</td>
<td>GIS(^a) operationalization of walkshed</td>
<td>Measurement criteria reported</td>
</tr>
<tr>
<td>Year of publication</td>
<td>Distance or time parameter</td>
<td>Criteria used with persons living with dementia</td>
</tr>
<tr>
<td>Journal name</td>
<td>Data sources and types</td>
<td>Criteria used with older adults</td>
</tr>
<tr>
<td>Journal discipline (if applicable)</td>
<td>GIS routines (if reported)</td>
<td>Measurement indicators reported</td>
</tr>
<tr>
<td>Country of lead author’s institution</td>
<td>Population focus</td>
<td>GIS based indicators</td>
</tr>
<tr>
<td>Study method</td>
<td>N/A(^b)</td>
<td>Data sources for indicator calculation</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>Method for indicator measurement or representation</td>
</tr>
</tbody>
</table>

\(^a\)GIS: Geographic Information System.

\(^b\)N/A: not applicable.

Beyond tracking the breadth (eg, diversity of methods) and location (eg, countries of origin) of literature, descriptive numerical summaries will examine 2 key topics. First, we will document the tools, data, and parameters used to define a walkshed. The review will make a contribution to the existing literature by documenting implementation approaches specific to the context of persons living with dementia. We will also compare these approaches to those used in studies of an older adult population. Second, we will chart criteria and indicators used to measure aspects of dementia-friendly neighborhood and community environments. By documenting indicators that scholars have operationalized using GIS-based analyses, we will make a key contribution to the transfer of the methodology.

The final scoping review will use descriptive results (eg, diversity of methods) represented using a combination of summary tables and figures (eg, Sankey diagrams). Limited textual information will support these visual elements. We will represent comparative results related to criteria and indicators as a larger data matrix. This matrix will visualize how
researchers have operationalized indicators in GIS for the 2 populations of interest. A longer textual description will contextualize these results. Finally, using thematic analysis, we will convey synthesized themes that capture nuance lacking in the descriptive and comparative results [41-43]. We expect to highlight considerations for the use of walkshed methodology not yet documented in recent studies focused on older adults [2,21]. We also expect to identify where criteria used to assess walkability for persons living with dementia and older adults converge and diverge. The risk of bias will not be assessed. This is consistent with the broad nature of our review question and the norms identified in the development of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [44,45].

**Results**

The results of the study and the submission of a manuscript for peer review are expected in June 2024.

**Discussion**

**Overview**

Scholars from the fields of planning, public health, urban design, gerontology, and architecture have produced a wealth of evidence and guidance related to walkability. Branching out from the “general population,” studies increasingly focus on targeted population groups. These foci better recognize the social, cultural, and demographic barriers and enablers to walking that shape one’s experience of the neighborhood. The proposed scoping review will synthesize the growing evidence base with specific reference to persons living with dementia. By including relevant studies focused on an older adult population, the review will also identify where current best practice for monitoring and evaluation diverges and converges for these populations. Expected findings include a detailed breakdown of current parameters and routines used to conduct walkshed analysis. Findings will also convey criteria that can be operationalized in GIS as indicators to assess barriers and facilitators to walking in a neighborhood setting.

Studies already identified here have documented monitoring and evaluation methods relevant to walkability for persons living with dementia. Methods include interviews, community survey techniques, and field audits of the neighborhood environment. Our planned focus on GIS-based walkshed analysis will further document a highly scalable monitoring and evaluation tool and technique.

**Limitations**

The final scoping review will be subject to limitations, despite following accepted methodological practice [39,40]. First, as a scoping review, there will not be a quality assessment of studies, which presents a risk of bias. Second, only English studies will be included, which will overemphasize evidence and practice from western countries. Third, we expect that the use of walkshed analysis for persons living with dementia will be an offshoot of techniques and literature focused on older adults. There may therefore be limited literature specific to persons living with dementia. To mitigate the risk of making assumptions about the transfer of methodological guidance from one population to another, we will explicitly track and compare findings across groups.

**Conclusions**

As far as we are aware, the proposed scoping review will be the first to provide comprehensive methodological or technical guidance for conducting walkshed analysis specific to persons living with dementia. There are 3 target audiences for this scoping review. These include applied academic researchers in the field of public health, applied academic researchers in the fields of urban planning and design, and evidence-based practitioners across these fields. Scholars identify neighborhood environments as an upstream source of barriers and enablers that shape walking behavior and associated health and well-being cobenefits [12,17,25]. Understanding the individual and population health impacts of neighborhood environments requires the expertise of health researchers and practitioners. Understanding how neighborhood environments came to be and how to reshape them through land-use and built-form interventions requires the expertise of planners and designers. By clearly documenting methods used in walkshed analysis, our goal is to spur increased collaboration across these disciplines to enable an evidence-informed approach to improving neighborhood environments for persons living with dementia.

**Acknowledgments**

The authors would like to thank the DemSCAPE team for their efforts.

**Data Availability**

The data generated and analyzed during this study will include content extracted from published, peer-reviewed journal articles. Full details about parameters, data sets, and Geographic Information System routines used in walkshed analysis, as well as a full list of associated indicators, will be reported in the scoping review publication. Additional data generated and analyzed during the study will be available from the corresponding author upon reasonable request.

**Conflicts of Interest**

None declared.
References


Abbreviations

GIS: Geographic Information System
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
**PRISMA-ScR:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

**Please cite as:**
Groulx M, Freeman S, Gourlay K, Hemingway D, Rossnagel E, Chaudhury H, Nouri M
JMIR Res Protoc 2024;13:e50548
URL: https://www.researchprotocols.org/2024/1/e50548
doi:10.2196/50548
PMID:38170573

©Mark Groulx, Shannon Freeman, Keone Gourlay, Dawn Hemingway, Emma Rossnagel, Habib Chaudhury, Mohammadjavad Nouri. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 03.01.2024. 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.
Overview of Retention Strategies for Medical Doctors in Low- and Middle-Income Countries and Their Effectiveness: Protocol for a Scoping Review

Norehan Jinah1*, MD; Ili Abdullah Sharin1*, MD; Pangie Bakit1*, MBBS; Izzuan Khirman Adnan1*, MBBS; Kun Yun Lee1*, MBBCBAO, MPH, DrPH
Centre of Leadership & Professional Development, Institute for Health Management, National Institutes of Health Malaysia, Shah Alam, Malaysia
* all authors contributed equally

Abstract
Background: The global shortage and maldistribution of health care workers, especially medical doctors, pose a significant threat to achieving the United Nations’ sustainable development goal 3 of ensuring well-being and healthy lives for all. Low- and middle-income countries (LMICs) are disproportionately affected by this crisis, with a high rate of brain drain from rural to urban areas, as well as to high-income countries. Various retention strategies have been implemented in different settings and organizations. However, their effectiveness remains underexplored, particularly in LMICs.

Objective: We aim to review the available retention strategies for medical doctors in LMICs and to determine the effectiveness of the various strategies. This review aims to compile relevant research findings on this issue to generate a thorough summary of all the retention strategies practiced in LMICs and, more importantly, to provide the current state of evidence of the effectiveness of these strategies in retaining medical doctors in countries with limited resources and high disease burden.

Methods: The structured framework given by Arksey and O’Malley will serve as the basis for conducting this scoping review. A comprehensive search strategy will be conducted across 4 electronic databases (PubMed, EBSCOHost, Scopus, and ScienceDirect). A systematic approach following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines will be executed to search, screen, review, and extract data from studies that meet predefined inclusion criteria. Data encompassing bibliographical information, study location, retention strategies, influencing factors, and outcomes (effectiveness) will be obtained from the selected studies using standardized data extraction. Endnote and Microsoft Excel will be used for reference management and removal of duplicate studies. A narrative synthesis will be performed after categorizing and analyzing all the extracted data to identify recurrent themes.

Results: This ongoing review will generate a comprehensive compilation of retention strategies implemented in LMICs to prevent brain drain among medical doctors. Data extraction is currently in progress, and completion is expected by early 2024. Themes regarding the types of strategies, influencing factors, and outcomes will be synthesized. The findings will highlight effective retention strategies, gaps, and challenges in implementation for the benefits of future research. By identifying common barriers and facilitators, this review will provide insights into enhancing the policies and initiatives for doctor retention in LMICs.

Conclusions: This scoping review explores the retention strategies practiced in LMICs and attempts to identify effective strategies from existing research. By evaluating the barriers and challenges that influence the effectiveness of these strategies, policymakers and health care leaders can strive to obtain balanced and optimal health human resources in their respective organizations and countries.
Introduction

Overview

Optimal functioning of the health care system relies heavily on the quantity and quality of health care workers (HCWs). The ability to enhance health service coverage and ensure that everyone in the population has access to the highest possible level of health relies on the availability, accessibility, acceptability, and quality of the health care workforce [1]. However, in recent years, a critical global challenge has emerged: a crisis characterized by a shortage and maldistribution of HCWs including doctors, nurses, and other professionals. This crisis poses a significant threat to achieving the United Nations Sustainable Development Goal 3 (promoting well-being and ensuring healthy lives for people of all ages) [2]. Nowhere is the impact of this challenge more acute than in low- and middle-income countries (LMICs), where limited financial and human resources compound the challenge of providing essential health care services. LMICs represent various nations in a diverse economic spectrum, encompassing low-income, lower-middle-income, and upper-middle-income countries. On the basis of the World Bank classification [3], low-income economies are defined as those with a gross national income (GNI) per capita of ≤US$1135 in 2022; lower-middle-income economies are those with a GNI per capita between US$1136 and US$4465; upper-middle-income economies are those with a GNI per capita between US$4466 and US$13,845. In contrast, high-income economies are those with a GNI per capita of ≥US$13,846. Remarkably, this classification encompasses 137 countries, constituting 63% of all nations worldwide [4].

The World Health Organization projected that by 2030, countries, predominantly LMICs, will face a substantial deficit of approximately 18 million health workers [5]. This significant shortage will greatly hamper the capacity to deliver vital health care services to those populations in most significant need. Simultaneously, countries with varied levels of socioeconomic development encounter different challenges in health human resource planning related to employment, deployment, and retention of their workforce [6]. Medical doctors are vital to the health care system because of their expertise, care, and impact. They play a crucial role in ensuring optimal health care delivery within health care institutions [7]. However, many parts of the world are grappling with a shortage of doctors, which stems from various factors such as emigration, imbalanced distribution between rural and urban areas, and shifts in population demographics [5]. There is a global shortage of approximately 2.8 million doctors [5], with LMICs bearing the brunt of this burden [8]. This scarcity is further exacerbated by the phenomenon of brain drain, with doctors from LMICs emigrating to high-income countries (HICs) due to better job offers and career progress. In some HICs, foreign-trained physicians sometimes amount to one-fifth of the total number of doctors in the workforce [9]. The movement of doctors from lower-to higher-income settings has resulted in substantial economic consequences, not solely due to the transfer of human capital, but more importantly, indirect impacts, such as increased morbidity and mortality associated with the loss of doctors [10].

Apart from brain drain to other countries, there is also a high rate of doctors’ resignations from the public health care system to join the more lucrative private sector, especially in countries with dual health care financing systems. Job dissatisfaction, including unsatisfactory work environment (lack of facilities, inflexible working hours, poor career progression, lack of professional autonomy, and ineffective management style) and unfavorable service conditions (poor salaries and funding, duplication of activities), are closely associated with high mobility, especially from the public to private sectors [11-13]. The phenomenon of HCWs resigning poses a significant obstacle to the advancement of the health care system in any given country, making it a topic of widespread concern [14]. The increasing number of resignations among HCWs, particularly in the Asia Pacific region, has been reported as the greatest threat to the development and sustainability of a resilient health care system in a recent study [15]. Despite efforts to increase supply and retain them, the workforce is still struggling to meet public health demands, as demonstrated in Spain and Brazil [16]. The same issue was also reported in India, where the vacancy rates were nearly 21% and 42% for medical officers and specialists at health centers, respectively [17].

Addressing the global health workforce crisis requires comprehensive strategies at both national and international levels. Retaining HCWs is a challenge in almost every country, be it HICs such as Canada, Australia, and Scotland or LMICs in Africa and Asia, especially in rural and remote areas [18,19]. Retention encompasses the duration between the initial engagement with a service and the eventual separation or departure from that service. It serves as a metric to gauge the length of time an individual stays within the service [20]. Retention strategies in the context of doctors encompass a range of interventions designed to attract and keep doctors in particular settings, such as remote or rural areas, with a specific focus on LMICs [21-23]. These strategies are aimed at mitigating doctor shortages and ensuring equitable health care access for underserved populations. Policy makers and health care managers must comprehend the factors that influence doctor retention and formulate targeted measures to address these factors [24,25]. Effective retention strategies contribute to the
stability and continuity of health care delivery, especially in regions with limited accessibility [21,23].

The significance of retention strategies lies in their capacity to yield various benefits, including cost savings, employee engagement, productivity, knowledge retention, competitive advantage, and organizational stability [26]. Addressing doctor shortages requires tailoring retention strategies to the unique challenges and requirements of health care professionals in each country. This is particularly critical in LMICs, where health care systems often contend with fragility, staffing shortages, limited resources, and a higher disease burden [27-29]. Furthermore, these countries grapple with brain drain challenges, issues of health care accessibility, weakened political will, and unstable governmental systems [30-33].

There are many known impediments to the retention of doctors, the most common being unfavorable working conditions, limited opportunities for career advancement, nonappealing financial incentive structures, unsupportive community environments, and the restriction of financial resources [34,35]. Other barriers include inadequate living standards, excessive workloads, insufficient equipment, lack of opportunities for skill enhancement and private practice, and unfair promotion practices [36]. In addition, stress, burnout, and insufficient work-life balance also play a role in doctors’ decision to leave [37]. Strategies aimed at addressing these barriers have been proposed and implemented at various levels and organizations, such as providing career development plans, ensuring minimum financial incentives, establishing avenues for private practice, enhancing work conditions, providing opportunities for skill improvement, and implementing transparent and equitable promotion systems.

Objective of Conducting the Scoping Review

Numerous publications have discussed the factors influencing the retention of doctors in LMICs [38-41], providing suggestions for various strategies and initiatives. However, there is limited research evaluating and summarizing the effectiveness of these strategies, particularly in LMICs. Therefore, the objective of this scoping review is to identify and delineate the available retention strategies for medical doctors in LMICs and to determine the effectiveness of these strategies.

To determine if prior research has addressed the same subject, we performed an initial exploratory literature review. Our search revealed the absence of existing or ongoing systematic reviews and scoping reviews related to our specific topic. McClain et al [42] primarily explored retention strategies and barriers concerning nurses, while Noya et al [43] concentrated on the rural and remote medical workforce, and Verma et al [22] focused on primary care doctors in general.

Conversely, our review aims to synthesize research evidence to generate an all-encompassing perspective on the effectiveness of retention strategies for doctors in LMICs. This synthesis will identify gaps in existing literature, pinpointing areas that require additional investigation within the context of doctor retention in resource-constrained countries with high disease burden. Our inclusive methodology considers a broad spectrum of studies and settings and delivers a comprehensive evaluation of these strategies.

Methods

Ethical Considerations

As the methodology for this scoping review solely entails reviewing and collecting data from existing literature without involving human participants, ethical clearance was waived by the Medical Research and Review Committee Malaysia.

Protocol Design

Overview

For this scoping review, we will use the methodological framework introduced by Arksey and O’Malley [44], who structured the review process into 5 stages. In addition, we enhanced the quality and rigor of our review based on the guidelines from the Joanna Briggs Institute Manual [45]. We will also incorporate the recommendations provided by Levac et al [46] to ensure consistency in assessing the studies during this scoping review. Transparent reporting will be ensured by using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [47]. We describe the protocol for this scoping review in five stages:

1. Formulating research questions
2. Identifying relevant studies
3. Selecting eligible studies
4. Charting the data
5. Collating, summarizing, and reporting the results

Stage 1: Formulating Research Questions

Following the recommendations given by Levac et al [46], we set our objective to explore strategies or interventions available for retaining doctors within health care institutions in LMICs and to identify effective measures to prevent doctor attrition. Therefore, we formulated two specific research questions for this review:

1. What are the retention strategies currently being implemented for doctors in LMICs?
2. Which strategies have been identified and evaluated as effective in retaining doctors in LMICs?

Stage 2: Identifying Relevant Studies

A meticulous search strategy plays a vital role in ensuring the inclusion of pertinent studies in scoping reviews. The research team has developed a comprehensive search strategy that encompasses various keywords and their synonyms related to the topic of interest. We selected search terms based on the research questions, including terms such as “retention,” “retain,” “maintain,” “doctor,” “physician,” and “general practitioner.” These terms have been used both individually and in combination following the iterative process inherent in the scoping review methodology.
The final search string, adhering to Boolean logic, takes the following form: (retention OR retain OR intention to leave OR intention to stay OR motivation to stay OR willingness to work) AND (doctor OR physician OR specialist OR general practitioner OR medical practitioner) AND (low- and middle-income countries OR LMIC). This meticulously designed search string aimed to gather all pertinent materials aligned with the objectives of this scoping review.

Various types of documents were screened during this stage, including journal articles, documents, or regulatory reviews, sourced from each of the 4 databases: PubMed, EBSCOHost, Scopus, and ScienceDirect. These databases were selected for their relevance to health and human resource services. During the screening process, if the available information in the title and abstract is insufficient to make an informed decision, the articles will be included for full-text screening. Adhering to the standard approach for conducting scoping reviews, we will not conduct quality appraisal of the included studies. An example of a preliminary MEDLINE (PubMed) search strategy is presented in Textbox 1.

Textbox 1. Example of MEDLINE (PubMed) search strategy.


Stage 3: Selecting Eligible Studies

The review process begins with the team convening to discuss decisions related to study inclusion and exclusion based on the principles of transparency, reproducibility, and rigor. This practice further advances a systematic and unbiased approach throughout the review process. The inclusion and exclusion criteria are presented in Textbox 2. We chose to focus primarily on studies published in English language due to their global prevalence, ensuring a comprehensive analysis, increased accessibility, and reduced language-related biases due to limited translation resources. Furthermore, the focus on studies published in English language streamlines the accessibility and application of research findings, making them readily available and applicable to a broader audience.

To maintain the scientific rigor of this review, we made a deliberate choice to exclude gray literature from our review, such as dissertations, essays, consensus, reports, theses, and government documents. However, while gray literature may provide valuable insights that supplement traditional academic
literature [48], it presents challenges in terms of systematic search and quality verification [49].

Following the PRISMA-ScR guidelines [47], the first step begins with identifying articles from various databases. Duplicates and irrelevant studies will then be removed. Abstracts or full texts will be evaluated based on predetermined inclusion and exclusion criteria to determine eligible studies. This screening process involves careful examination of both the retrieved search results and their reference lists. To ensure the most relevant search results, we will refine the literature search throughout the review process. At least 2 investigators will independently assess the eligibility of publications by reviewing their titles and abstracts. Publications deemed relevant to this scoping review are obtained in full text and reviewed against the same inclusion criteria.

In cases of disagreement during publication selection, both reviewers will revisit the full-text articles to reach a consensus. If consensus cannot be reached, an impartial third reviewer will be consulted to resolve the disagreement. Consistent meetings and discussions at different stages of the article review process are essential to maintain alignment, address challenges, refine search strategies, ensure consistency, and foster a collaborative and efficient approach. The scoping review will record and report reasons for excluding sources of evidence in the full text that do not meet the inclusion criteria. The reporting of the review will incorporate a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1), which visually presents the screening and selection process [50].

**Textbox 2.** Inclusion and exclusion criteria of the study selection process.

<table>
<thead>
<tr>
<th><strong>Inclusion criteria</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Publication year: January 2013 to February 2023</td>
</tr>
<tr>
<td>• Language in publication: English</td>
</tr>
<tr>
<td>• Research location: low- and middle-income countries (LMICs)</td>
</tr>
<tr>
<td>• Target population: medical doctors</td>
</tr>
<tr>
<td>• Types of documents: journal articles, documents, or regulatory reviews with proper references</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Exclusion criteria</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Publication year: before January 2013 and after February 2023</td>
</tr>
<tr>
<td>• Language in publication: other languages</td>
</tr>
<tr>
<td>• Research location: other than LMICs</td>
</tr>
<tr>
<td>• Target population: other health care professionals</td>
</tr>
<tr>
<td>• Types of documents: dissertations, essays, consensus, government documents, reports, and theses that do not have any proper references</td>
</tr>
</tbody>
</table>
Stage 4: Charting the Data

The data extracted from the full-text articles will be organized into a data extraction table using Microsoft Excel (Microsoft Corporation). The data table will be structured to accommodate the characteristics of the data. The aim of charting the data is to create a descriptive summary of the results to address the objectives of the scoping review and to answer the research questions. This process facilitates the categorization of information before proceeding with further tabulation. For reference, Textbox 3 presents the categories corresponding to each characteristic in the data extraction table. In an iterative process, investigators will continually gather data and keep the data extraction table up-to-date. If significant data are found in records initially not designated for extraction, the data extraction form will be revised, and these additional data will be retrieved from the records already reviewed.
**Textbox 3.** Preliminary data extraction table.

<table>
<thead>
<tr>
<th>Basic characteristics and description</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Bibliographical data</td>
</tr>
<tr>
<td>• First author and year of publication of the article</td>
</tr>
<tr>
<td>• Article title</td>
</tr>
<tr>
<td>• A succinct description of the content of the article</td>
</tr>
<tr>
<td>• Country</td>
</tr>
<tr>
<td>• Name of the low- and middle-income countries</td>
</tr>
<tr>
<td>• Aims or purpose of the study</td>
</tr>
<tr>
<td>• Expresses the intention or aspiration of the research</td>
</tr>
<tr>
<td>• Type of study</td>
</tr>
<tr>
<td>• Study design or methodology</td>
</tr>
<tr>
<td>• Which type of study was conducted?</td>
</tr>
<tr>
<td>• Study population</td>
</tr>
<tr>
<td>• Physician—specialty or department</td>
</tr>
<tr>
<td>• Number of people involved</td>
</tr>
<tr>
<td>• Inclusion and exclusion criteria of the study</td>
</tr>
<tr>
<td>• Demographic characteristics</td>
</tr>
<tr>
<td>• Other characteristics</td>
</tr>
<tr>
<td>• Study location</td>
</tr>
<tr>
<td>• Location characteristics (urban, rural, or remote or hospital or district, state, or area)</td>
</tr>
<tr>
<td>• Institution (name)</td>
</tr>
<tr>
<td>• Factors influencing retention</td>
</tr>
<tr>
<td>• Financial or career and professional or working conditions, personal, cultural, or living conditions factors</td>
</tr>
<tr>
<td>• Retention strategy</td>
</tr>
<tr>
<td>• Strategy type or focus (education and regulatory or ii. monetary compensation or iii. management, and environment and social support)</td>
</tr>
<tr>
<td>• Strategy name</td>
</tr>
<tr>
<td>• Strategy characteristics, content, and description</td>
</tr>
<tr>
<td>• Strategy implementation (levels, duration, and date)</td>
</tr>
<tr>
<td>• Outcomes measure</td>
</tr>
<tr>
<td>• Description of the result (effective or successful to retain)</td>
</tr>
<tr>
<td>• How were the turnover and results assessed?</td>
</tr>
<tr>
<td>• Barriers and challenges</td>
</tr>
<tr>
<td>• Barriers and challenges in implementing the strategies</td>
</tr>
<tr>
<td>• Study limitations</td>
</tr>
<tr>
<td>• Weaknesses within the research design that may influence the outcomes and conclusions of the research</td>
</tr>
</tbody>
</table>
**Stage 5: Collating, Summarizing, and Reporting the Results**

The primary goal of the scoping review is to present the narrative findings of existing literature through an analytical framework or thematic construction, without the requirement to assess the quality or significance of each study. We will use a traditional integrative review approach to compile all the identified materials. Our objective is to identify recurring themes across research and synthesize data from the selected studies. Using these themes as guidelines, we will create a literature map and present it in the form of a table, summarizing the publications and their respective characteristics.

The results of the scoping review will be organized into tables that categorize the characteristics of each publication. Accompanying these results will be narrative summaries that describe how each result relates to our research questions, including any unexpected or particularly notable findings. We will also address any gaps observed in the literature, research needs, and implications for practice. Subsequently, the outcomes of this review will be shared with relevant stakeholders, and their expertise and perspectives will be incorporated.

**Results**

This review will provide a comprehensive mapping of existing research and literature pertaining to the retention of medical doctors in LMICs to enhance the understanding of the complex dynamics of doctor retention. It will also assess the current knowledge and pinpoint any gaps in the literature, focusing on factors influencing doctor retention and effective retention strategies such as financial incentives, working conditions, career advancement opportunities, and personal motivations.

Furthermore, this review can offer insights into best practices and approaches for retaining doctors in LMICs to guide policy makers and health care administrators who struggle with retention challenges. They can customize the best policy recommendations based on specific needs and obstacles in local settings to improve doctor retention rates in their respective organizations and governments.

The review was initiated in May 2023, and the research protocol was finalized in July 2023. We registered the review with the Malaysian National Medical Research Register (NMRR ID-23-01994-OGW). The search, which was concluded in August 2023, yielded 9141 articles. The PRISMA flow diagram will be used to illustrate the flow of the literature search in this review [50]. The results will be presented using charts and tables, supplemented by a narrative description. Any existing literature gaps will be identified, and the significance of our findings will be emphasized in the subsequent discussion section. The review is expected to be concluded in January 2024, with the outcomes published in a journal for wider dissemination.

**Discussion**

**Overview**

Adequate investment in health care capacity is imperative to move toward the United Nations’ sustainable development goals, specifically goal 3 (ensuring good health and well-being) and goal 10 (reducing inequalities), and to achieve various global development objectives, with a robust health care workforce being the top priority. Therefore, establishing a comprehensive plan that encompasses effective retention strategies to complement medical education reforms is vital to cultivating a health care environment that is equitable and resilient at both regional and global levels. Our focus on retention strategies for medical doctors is driven by their unique challenges and critical roles in health care. Doctors hold central positions in health care delivery, not only providing medical expertise but also taking a leadership role in influencing critical patient care decision-making, and their turnover can have significant negative impacts on patient care and quality of health care services [51]. Furthermore, doctors are the most affected by the brain drain crisis, especially in LMICs, leading to a significant financial burden and experience loss. Therefore, prioritizing doctor retention is vital for mitigating brain drain, reducing productivity and financial loss, and sustaining effective health care service delivery.

The shortage of doctors in LMICs represents a pressing concern that demands immediate attention and concerted efforts on a global scale, in view of its significant impact on public health. This predicament has a direct adverse effect on the health and welfare of populations residing in LMICs, as doctor shortages can impede access to crucial medical services, ultimately resulting in preventable illnesses and higher mortality rates. Moreover, cross-border brain drain exacerbates existing health care inequalities both within and between countries. Persistent disparities in the accessibility of health care services, if they continue to exist, will disproportionately affect rural and underserved areas with limited resources, thereby perpetuating social and economic inequalities and impeding advancements toward achieving universal health care coverage.

**Expected Outcomes**

This scoping review will present a comprehensive overview of retention strategies that have been proposed, practiced, and evaluated in LMICs as a response to overcome the challenges faced in retaining medical doctors and preventing brain drain. These strategies may encompass a wide array of approaches, including financial incentives, opportunities for professional development, initiatives to promote work-life balance, and support for career advancement. Moreover, the focus on LMICs may shed light on distinct regional or country-specific challenges and variations in customized strategies. It may also highlight the varied effectiveness of different strategies, depending on the contextual factors at play. It is unlikely to be a one-size-fits-all solution, as certain strategies may exhibit promising outcomes in bolstering medical doctor retention, while others may demonstrate limited impact depending on the local settings.
In short, this review will present common barriers and facilitators that significantly influence the successful implementation of retention strategies for doctors in LMICs. By exploring the challenges encountered during strategy implementation, we also aim to offer a more comprehensive and nuanced understanding of the factors influencing the effectiveness of doctor retention strategies in LMICs. This, in turn, can contribute to improving the retention of medical doctors in LMICs, aligning with the Sustainable Development Goal 3 goal of promoting well-being and ensuring healthy lives for everyone. Comprehension of these elements has the potential to aid policy makers and health care administrators in developing more relevant interventions and prioritizing effective strategies.

Since it is likely that different contexts play a critical role in the outcomes of various retention strategies, we will also attempt to address this connection in our review. Certain strategies, if proven successful, can also be modified and embedded within a broader health care ecosystem to benefit a wider group of health care professionals. Common factors contributing to brain drain among HCWs include financial rewards, career development, hospital infrastructure, political issues, and family issues [52]. While we focus on the dynamics surrounding medical doctors and the customized retention approach for them in this review, as the challenges faced by doctors may be unique and differ significantly from those of other groups, the comparison and extrapolation of various retention strategies for different health care professionals is a worthy topic for future research or review.

Review Limitations
This review has several limitations that deserve further discussion. First, the language restriction used in the search strategy may have unintentionally excluded relevant studies published in languages other than English. This is a significant concern because many LMICs have diverse linguistic landscapes with numerous languages. The decision to focus primarily on English was necessitated by practical considerations, such as the broader availability and accessibility of English-language research. Furthermore, we believe that the exclusion of non–English-language studies would minimize language-related biases in the review process, given the limited access to translation resources in our setting.

Another limitation of this review is the exclusion of gray literature. This decision is influenced by the difficulties associated with accessing gray literature, which encompasses issues of limited availability, inconsistent indexing, variable accessibility, and challenges in assessing the quality and reliability of information. By excluding gray literature, there is a risk of missing important findings and diverse perspectives not found in peer-reviewed academic sources. Nevertheless, although gray literature can offer valuable insights as a complement to conventional academic literature [48], it introduces difficulties in systematic retrieval and quality assessment [49], thus making it difficult to maintain the scientific rigor of this review.

In addition, this review is likely to include studies with different levels of methodological rigor and quality, and this could potentially affect the overall reliability of its conclusions and may introduce heterogeneity into our analysis. Moreover, this review aims to provide a comprehensive overview of existing literature regarding effective retention strategies for doctors in LMICs; thus, the analysis results are likely to be less in depth compared with systematic reviews that follow a more rigorous and narrowly focused methodology. Nonetheless, this broad approach is valuable for summarizing the diversity of strategies and findings in the field of doctor retention in LMICs, allowing for a holistic understanding of the subject.

Conclusions
This scoping review is fundamental in providing a better understanding of the practical implications of various retention strategies for doctors in LMICs and in drawing valuable lessons from effective strategies in existing literature. Furthermore, by highlighting emerging trends and identifying implementation challenges within LMICs, this review will pave the way for more precisely targeted policies and interventions to strengthen doctor retention in the most needed regions. It also offers valuable guidance to policy makers and health care administrators by showcasing best practices with positive outcomes, thereby refining their approach to addressing attrition and brain drain.

Acknowledgments
The authors would like to thank the Director General of Health, Malaysia for the permission to publish this paper.

Data Availability
Data sharing is not applicable to this paper, as no data sets were generated or analyzed in this review.

Conflicts of Interest
None declared.

References


6. Health workforce. World Health Organization. URL: https://www.who.int/health-topics/health-workforce#tab=tab_1 [accessed 2023-08-12]


18. Increasing access to health workers in remote and rural areas through improved retention: global policy recommendations. World Health Organization. 2010. URL: https://iris.who.int/handle/10665/44369 [accessed 2023-08-12]


50. PRISMA flow diagram. PRISMA. URL: http://prisma-statement.org/prismastatement/flowdiagram.aspx [accessed 2023-08-12]


Abbreviations

GNI: gross national income
HCW: health care worker
HICs: high-income countries
LMICs: low- and middle-income countries
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

©Norehan Jinah, Ili Abdullah Sharin, Pangie Bakit, Izzuan Khirman Adnan, Kun Yun Lee. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.01.2024. 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

Seasonal Malaria Chemoprevention Therapy in Children Up To 9 Years of Age: Protocol for a Cluster-Randomized Trial Study

Mahamoudou Toure1, MSPH, MD, PhD; Jeffrey G Shaffer2, PhD; Daouda Sanogo1, MPH, MD; Soumba Keïta1, MSc, PharmD; Moussa Keïta1, MSc, PhD; Fousseyni Kane1, MSc, MD; Bourama Traore1, PharmD, MSc; Djeneba Dabitao1, PharmD, PhD; Aissata Kone3, MPH, MD; Cheick Oumar Doumbia1, MPH, MD; Joseph Keating2, PhD; Joshua Yukich2, PhD; Helle H Hansson4, PhD; Alyssa E Barry5,6, PhD; Mahamadou Diakité1, PharmD, DPhil; Michael Alifrangis4, PhD; Seydou Doumbia1, MD, PhD

1University Clinical Research Center, Université des Sciences, des Techniques et des Technologies, Bamako, Mali
2Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, United States
3Mali National Malaria Control Program, Bamako, Mali
4Center for Medical Parasitology, Department of Immunology and Microbiology, University of Copenhagen, Copenhagen, Denmark
5Institute for Mental and Physical Health and Clinical Translation (IMPACT) and School of Medicine, Deakin University, Geelong, Melbourne, Australia
6Life Sciences Discipline, Burnet Institute, Melbourne, Australia

Corresponding Author:
Mahamoudou Toure, MSPH, MD, PhD
University Clinical Research Center
Université des Sciences, des Techniques et des Technologies
Point G
Bamako
Mali
Phone: 223 66778912
Email: mah.toure@gmail.com

Abstract

Background: Seasonal malaria chemoprevention (SMC) is recommended by the World Health Organization for the sub-Sahel region in sub-Saharan Africa for preventing malaria in children 3 months old to younger than 5 years. Since 2016, the Malian National Malaria Control Program has deployed SMC countrywide during its high malaria transmission season at a rate of 4 monthly cycles annually. The standard SMC regimen includes sulfadoxine-pyrimethamine (SP) plus amodiaquine (AQ). Resistance against SP is suspected to be rising across West Africa; therefore, assessing the effectiveness of an alternative antimalarial drug for SMC is needed to provide a second-line regimen when it is ultimately needed. It is not well understood whether SMC effectively prevents malaria in children aged 5 years or older.

Objective: The primary goal of the study is to compare 2 SMC regimens (SP-AQ and dihydroartemisinin-piperaquine [DHA-PQ]) in preventing uncomplicated Plasmodium falciparum malaria in children up to 9 years old. Secondly, we will assess the possible use of DHA-PQ as an alternative SMC drug in areas where resistance to SP or AQ may increase following intensive use.

Methods: The study design is a 3-arm cluster-randomized design comparing the SP-AQ and DHA-PQ arms in 2 age groups (younger than 5 years and 5-9 years) and a control group for children aged 5-9 years. Standard SMC (SP-AQ) for children younger than 5 years was provided to the control arm, while SMC with SP-AQ was delivered to children aged 3 months to 9 years (arm 2), and SMC with DHA-PQ will be implemented in study arm 3 for children up to 9 years of age. The study was performed in Mali’s Koulikoro District, a rural area in southwest Mali with historically high malaria transmission rates. The study’s primary outcome is P falciparum incidence for 2 SMC regimens in children up to 9 years of age. Should DHA-PQ provide an acceptable alternative to SP-AQ, a plausible second-line prevention option would be available in the event of SP resistance or drug supply shortages. A significant byproduct of this effort included bolstering district health information systems for rapid identification of severe malaria cases.

Results: The study began on July 1, 2019. Through November 2022, a total of 4556 children 3 months old to younger than 5 years were enrolled. Data collection ended in spring 2023, and the findings are expected to be published later in early 2024.
Conclusions: Routine evaluation of antimalarial drugs is needed to establish appropriate SMC age targets. The study goals here may impact public health policy and provide alternative therapies in the event of drug shortages or resistance.

Trial Registration: ClinicalTrials.gov NCT04149106, https://clinicaltrials.gov/ct2/show/NCT04149106

International Registered Report Identifier (IRRID): DERR1-10.2196/51660

(JMIR Res Protoc 2024;13:e51660) doi:10.2196/51660

KEYWORDS
malaria; seasonal malaria chemoprevention; RCT; randomized; controlled trial; controlled trials; parasite; parasites; mosquito; mosquitoes; vector-borne; malignant; antimalarial; age; Plasmodium falciparum, protocol, cluster-randomized trial; child; children; infant; infants; pediatric; pediatrics; clinical trial; clinical trials; drug; drugs; pharmacy; pharmacology; pharmaceutical; pharmacies; pharmaceuticals; pharmaceutical; medication; medications

Introduction

Malaria is endemic to south and central Mali, where over 90% of its approximately 17.6 million population is at risk for infection [1]. The disease primarily burdens rural areas that maintain suitable larval habitats and lack access to adequate health care [1]. Malaria transmission is highly seasonal in Mali (length of the transmission periods varies from 3 to 6 months), with a peak of malaria cases at the end of the rainy season (October through November), though it may be affected by irrigation schemes [2,3]. Deaths due to malaria registered at health centers totaled 1050 in 2017, with 669 (63.7%) occurring among children younger than 5 years. However, these results are likely substantially underreported. For instance, in 2015, health facilities in Mali reported over 2.3 million confirmed malaria cases and 1544 malaria deaths, but actual estimates were 7.5 million and 21,000, respectively [1].

Since 2007, support from the US Presidential Malaria Initiative program and other sponsors resulted in a 50% reduction in Mali’s malaria burden. These efforts have been carried out through increased preventive and treatment measures such as long-lasting insecticide-treated mosquito nets (LLINs), indoor residual spraying, artemisinin-based combination therapies (ACTs), and intermittent preventive treatment of pregnant women (IPTp); and more recently, seasonal malaria chemoprevention (SMC). The most widely implemented malaria control interventions are the joint use of LLINs and rapid treatment of malaria cases with ACTs, IPTp, and SMC. Indoor residual spraying was previously implemented on a small scale in several Malian districts but is currently implemented in only a single district in central Mali.

Despite the broad deployment of these interventions, malaria prevalence and incidence rates remain high in Mali according to the routine Malaria Indicators Surveys [4]. Since 2010, the International Centers of Excellence in Malaria Research (ICEMR), in collaboration with Mali’s National Malaria Control Program (NMCP), has identified significant constraints to malaria control implementation strategies in Mali, including the upward shift of the prevalence of infection and incidence of disease in children younger than 5 years to children aged 5-9 and 10-14 years [5,6]. These findings are particularly significant as the World Health Organization (WHO) guidelines recommend SMC only for children 3 months old to younger than 5 years. Therefore, the Mali NMCP and its partners have expressed SMC effectiveness and intervention strategies as key research priorities.

While the sulfadoxin and pyrimethamine regimen (SP) remains effective in West Africa, resistance to this regimen has been observed in East African regions where Plasmodium falciparum is highly prevalent. SMC is not currently recommended for countries in southern and eastern Africa due to widespread resistance, although there are some locations where transmission patterns suggest potential suitability [7,8]. In western parts of Africa, higher frequencies of the triple dihydrofolate reductase mutants and the quadruple mutant (triple dihydrofolate reductase plus dihydropyrimidinase synthetase 437) associated with significant resistance to SP have been observed in children receiving SMC in both Mali and Senegal [9,10]. Recent studies have also reported SP resistance in several parts of Mali [11,12]. The wide-scale deployment of SMC in sub-Saharan African countries required increased focus on P falciparum resistance as well as ongoing assessments of new alternative drug combinations for SMC as this strategy has proven to be effective in reducing the impact on severe malaria and mortality in children 3 months old to younger than 5 years [13]. A recent study in Burkina Faso suggests that higher dosages and extended dosing of dihydroartemisinin-piperaquine (DHA-PQ) to 4 monthly doses (and cover the entire high malaria transmission period) could reduce malaria incidence up to 58% during the peak transmission season [14,15]. DHA-PQ has demonstrated excellent efficacy for chemoprevention and benefits from the long half-life of piperaquine and offers a protective efficacy of 98% against malaria in Thai adults when administered monthly [7]. In Senegalese children, similar monthly malaria incidence was observed among children receiving monthly DHA-PQ vs SP-amodiaquine (AQ) as SMC regimen [8].

A study in Uganda reported a 58% greater protective efficacy for DHA-PQ vs SP-AQ based on monthly administration among children younger than 5 years [9]. This study aims to determine whether SMC effectively prevents malaria in children aged 5-9 years. The evidence-based approaches used will guide policy in Mali and other countries in West Africa using SMC.

Methods

Ethical Considerations

The study was approved by the University of Sciences, Techniques and Technologies Ethics Board under the following reference (N°2019/04/CE/FMPOS). The trial will report the
efficacy of 2 SMC subtypes (SP-AQ and DHA-PQ) for preventing *P. falciparum* malaria in children aged 5-9 years in Mali. Should the findings show that SMC is efficacious for children in this age group, it may impact policy SMC delivery in Mali.

**Study Site**

Mali’s Koulikoro District is situated in its southern region, approximately 50 miles (18 km) north of Bamako and 255 miles (410 km) from the Guinea border [11]. The district maintains 21 health zones and 71 community health posts. The current population in the district is 282,570, with approximately 4% of its population <1 year of age and 18% between 1 and 4 years of age. The total number of villages a community health center covers ranges from 8 to 31. The most populous village is Kolebougou, with 34,712 persons; the least populated is Souban Village, with 5085 persons. Ongoing malaria control activities include case management (rapid diagnostic tests and ACTs), IPTp, SMC, and LLIN use. District health centers maintain clinical and laboratory research capacity and full-time staff and clinicians for malaria screening and patient care. The site was chosen for the proposed research because of its diverse range of malaria control interventions, collaborative research agreements with the University of Sciences, Techniques and Technologies of Bamako, Mali (USTTB), high malaria transmission rates, and rural location.

**Study Design and Population**

The study was as a 3-arm cluster-randomized design, as illustrated in Figure 1.

**Figure 1.** Study design of a 3-arm cluster-randomized trial on age targets for season malaria chemoprevention. The study was performed from July 2019 to March 2023. AQ: amodiaquine; DHA: dihydroartemisinin; PQ: piperaquine; SP: sulfadoxine-pyrimethamine.

The study population includes children aged 3 months to 9 years in each village. Each study arm includes 3 corresponding public health units (PHUs) that differ with respect to their ecology, proximity to the Niger River, and region in the context of Koulikoro District (northern, central, and southern).

**Cluster Selection**

A total of 9 villages were randomly selected from Mali’s Koulikoro District for participation. The selection strategy focused on 3 aspects: proximity to the Niger River (which lies in the southern part of the region) and the central and northern parts of the sampling region. Within these strata, villages were rank-ordered according to their populations. Random selection was carried out such that villages of high, medium, and low populations were sampled from each of these strata. A total of 3 villages (1 village per stratum with different ranking in terms of population size) were assigned to a single treatment arm (Figure 2).
Figure 2. Map of the 9 study sites in Mali’s Koulikoro District. The 9 study sites are near the central part of Mali’s Koulikoro District. Study sites were selected as 3 villages near the Niger River, 3 villages in the northern part of Koulikoro District, and 3 villages in the central part of Koulikoro. AQ: amodiaquine; DHA: dihydroartemisinin; PQ: piperaquine; SP: sulfadoxine-pyrimethamine.

The southernmost villages (Gouni, Kenenkoun, Kamani) are situated near the Niger River, while the remaining villages (Doumba, Sinzani, Koula; and Sirakorola, Monzombala, and Chola) lie in the northern and southern parts of the region above the Niger River, respectively.

Randomization
Probability proportional to population size sampling was used for allocating villages to the study arm. Random treatment assignment was balanced according to population size in each stratum (low, medium, and high), resulting in a 3 by 3 Latin square arrangement (Table 1).

Table 1. Study site classification per sampling region within the Koulikoro District, Mali.

<table>
<thead>
<tr>
<th>Region</th>
<th>Population size</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>South</td>
<td>Small</td>
<td>Kamani (C)</td>
<td>Kenenkoun (B)</td>
</tr>
<tr>
<td></td>
<td>Gouni (A)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td></td>
<td>Doumba (A)</td>
<td>Koula (C)</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sinzani (B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>Large</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chola (C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monzombala (B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sirakorola (A)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A, B, and C denote the 3 study arms.

Specifically, PHUs were rank-ordered according to their total populations, and the top 3 PHUs (in terms of their populations) were chosen to represent Koulikoro’s southern, central, and northern regions. The PHUs were then randomized to treatment according to the Latin square arrangement shown in Table 1.

Eligibility and Enrollment
Participants were eligible to participate if they were residents of the sampled village, aged 3 months to 9 years of age at enrollment, were asymptomatic of current chronic diseases, and did not have a history of allergies to SP, AQ, or DHA-PQ therapies. Informed, written consent is required for all study participants annually. Consent was administered in oral or written formats and included a full description of voluntary participation, the right to withdraw from the study at any time, and the right to refuse to answer any question or participate in any research component.

Study Outcomes
The primary study outcome is the incidence of severe or uncomplicated *P falciparum* malaria. The population denominator for incidence calculations was derived from the house-to-house enumeration with household member listing. Secondary outcomes were abstracted from censuses and enumeration lists for selected villages, including village-level distributions on populations, age, gender, residential status, and LLIN use. Exhaustive selection was carried out in villages for selected PHUs for those children who met the eligibility criteria.
and for whom consent of parents or tutors for enrollment was obtained. A baseline malaria infection prevalence survey was carried out before the initial SMC campaign. Follow-up surveys at community health centers were also carried out to assess malaria incidence among study participants. Monthly SMC administration and compliance with treatment assessment were captured via reports by caregivers and measurement of AQ metabolite by enzyme-linked immunosorbent assays. Demographic data, including age, sex, and clinical parameters (including temperature, pulse, and respiration rate), were recorded at a community health center visit. Malaria cases were defined as fever or a history of fever within the past 48 hours associated with either positive malaria rapid diagnostic test or a positive blood smear prepared during that visit.

**Hypotheses and Rationale**

The primary hypotheses were as follows:

1. Malaria incidence in children was at least 10% higher in the 5- to 9-year age group without SP-AQ than in children in the 5- to 9-year age group with SP-AQ.
2. Malaria incidence in children 3 months old to younger than 5 years receiving DHA-PQ was not statistically different from children in the same age group receiving SP-AQ.

**Data Collection and Management**

Confirmed malaria case incidence data will be collected from public health facilities through electronic data capture. A REDCap (Research Electronic Data Capture; Vanderbilt University) database and mobile app were used for this study, and data synchronization will be performed daily for quality checking complemented with data queries distributed to the centralized data center at USTTB.

**A Priori Sample Size and Statistical Power**

The power and sample size assessment is based on the ability to detect a clinically significant difference in malaria incidence proportions between the 3 comparison groups in year 1. The minimum clinically meaningful differences between the comparison groups was set at 10%. The sample sizes were calculated assuming a design effect of 2 with a 2-sided type I error set at 0.05 and power set at 80%. The number of required subjects per study arm for detecting at least 10% differences in malaria incidence proportions was 1552 in year 1. This result was inflated by 233 subjects in year 2 to account for new subjects due to increased child ages, yielding 1785 participants per arm total or 5355 participants overall.

**Phase 2 (Years 3-5, 2021-2023)**

This phase involved expansion to 71 public health posts in years 3-5 (Figure 3). The district-wide data collection plan will build on the efforts for years 1-2 and focus on (1) cluster-randomized health posts to the 3 treatment arms and (2) semiannual community cross-sectional surveys. More specifically, the 71 health posts over the entire district were divided into 4 regions (Table 2).

**Figure 3.** Map of the selection approach for 71 public health posts in Mali’s Koulikoro District.
Within the stratum, health posts will be randomized to 1 of 3 treatment arms (SP-AQ <5 years, SP-AQ <10 years, and DHA-PQ <10 years). Allocation was proportional to population size based on census populations up to 9 years of age while balancing on population sizes across the 3 study arms. For regions maintaining a total number of health posts that were not divisible by 3, 1 or 2 adjacent health posts were considered as a single unit to ensure divisibility (Figure 3).

The health facilities selected in years 1 and 2 were also selected in years 3-5 and will leverage the process and training already in place in 9 health centers. Patterns over the entire 5-year study period will be analyzed for these 9 villages. Semiannual cross-sectional household surveys will permit assessment of community-level effects over the entire 5-year period.

### Results

The study began on July 1, 2019. Through November 2022, a total of 4556 children were enrolled during the pilot phase (2019-2020) in 9 villages across the Koulikoro Health District. In 2022, preliminary findings have been presented at the American Society of Tropical Medicine Conference [16] and published in the *American Journal of Tropical Medicine and Hygiene* [17]. SP-AQ and DHA-PQ were highly effective in reducing *P. falciparum* malaria in children 5-9 years in Koulikoro, Mali, at both the pilot and district-wide study phases. Data collection ended in spring 2023, and the findings are expected to be published later in early 2024. Results will be summarized and reported using both intention-to-treat and per-protocol analyses.

### Discussion

This study was designed to assess the effectiveness and efficacy of SP-AQ in children aged up to 9 years and of DHA-PQ as an alternative in the event of drug resistance to AQ or SP in West Africa and, more broadly, across Africa. The trial included 2 SMC regimens to provide a viable alternative therapy in the event of drug resistance or shortages in drug supply. The study enrolled over 5000 subjects in 9 villages in Mali’s Koulikoro District. The expanded phase of the study covered the entire Koulikoro District. The initial phase of the study covered 9 villages, which may be considered as a pilot for the larger district-wide trial.

This study provides an opportunity to directly measure the effect of the 2 drugs while extending SMC to older children in response to recent reports showing an age shift in malaria incidence and prevalence among older children. The study site here comprised different ecological settings represented countrywide with differential malaria transmission intensities. For instance, each study arm was composed of a village located along the Niger River where longer transmission seasons have historically occurred, a village located in the central part of Koulikoro District where malaria transmission peaks between July and October, and a village in the southern part of Koulikoro District where malaria transmission season has been historically shorter (August to November) than the other 2 locations. Also, all selected villages maintain community health facilities where malaria case management is done routinely. Additionally, a routine assessment of treatment compliance by measuring amodiaquine metabolite in the venous blood of children after treatment will be performed for the first time to determine SMC compliance across all 4 dosing periods. While the number of villages was initially restricted to 3 per study arm, all 71 PHUs in Koulikoro will be covered across the 3 study arms should the study meet the go/no go criteria for continued funding.

Particular challenges at the outset of the study planning involved garnering support from regional health administrators about using DHA-PQ instead of SMC and extending SMC use to children aged 5-9 years. Since the inception of Mali’s NMCP, a key partner in this effort made apparent the study’s rationale in terms of its implications on potential drug resistance with health agents, community leaders, and health workers before the SMC campaign was launched in Mali. Determining the appropriate age target for preventive therapies remains challenging for developing countries like Mali. The age captured during the most recent census was used to estimate the number of children failing in the <5 and 5-9-year age groups, which was usually reported by the participants’ parents or guardians. However, approximately 2% of subjects were misclassified during the SMC administration due to erroneous reporting during the census. Therefore, more accurate estimations of participant age (which was used to define the comparison groups) were captured for the study here at baseline.

SMC covers the rainy season that coincidentally occurs during school vacations and agricultural activities in rural Mali. Therefore, a child’s movement during that period is often a cause for losses to follow-up as children frequently might travel to Mali’s capital city of Bamako from July to September, while others will move with their family to the farms (hamlets) until October when school commences.

We believe that this study provides a somewhat novel study design that will aid researchers in assessing age targets in locations where controlled studies are not feasible.
Acknowledgments

We are grateful to the National Institutes of Health for funding this study. We also thank the kind community leaders and residents in Koulikoro District for their guidance, support, and participation. We thank Dr Donald J. Krogstad for his lifelong mission for improving health outcomes in West Africa and whose efforts laid the groundwork for this study. We dedicate the study to those impacted by malaria and hope the study contributes to reducing the burden of malaria in Koulikoro and across Mali. This work was supported by Cooperative Agreements from the National Institutes of Allergy and Infectious Diseases (NIAID) for the West African International Center of Excellence for Malaria Research (ICEMR): U19AI089696 and U19AI129387 (from 2010 to 2017, 2017 to 2024, and 2016 to 2021, respectively). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

Data will be archived on the ClinEpiDB institutional repository [18] to ensure long-term availability, preservation, and access. The data sets generated during this study are available from the corresponding author on reasonable request as well as from the West Africa International Centers of Excellence in Malaria Research (ICEMR) Program Director by e-mail (sdoumbi@icermali.org or sdoumbia@gmail.com)

Conflicts of Interest

None declared.

Multimedia Appendix 1

Peer review report from the National Institute of Allergy and Infectious Diseases Special Emphasis Panel Limited Competition: Revision Applications for International Centers of Excellence for Malaria Research (NIH, USA).

References


Abbreviations

\textbf{ACT:} artemisinin-based combination therapy \\
\textbf{AQ:} amodiaquine \\
\textbf{DHA:} dihydroartemisinin \\
\textbf{ICEMR:} International Centers of Excellence in Malaria Research \\
\textbf{IPTp:} intermittent preventive treatment of pregnant women \\
\textbf{LLIN:} long-lasting insecticide-treated mosquito net \\
\textbf{NMCP:} National Malaria Control Program \\
\textbf{PHU:} public health unit \\
\textbf{PQ:} piperaquine \\
\textbf{REDCap:} Research Electronic Data Capture \\
\textbf{SMC:} seasonal malaria chemoprevention \\
\textbf{SP:} sulfadoxine-pyrimethamine \\
\textbf{USTTB:} University of Sciences, Techniques and Technologies of Bamako, Mali \\
\textbf{WHO:} World Health Organization

©Mahamoudou Toure, Jeffrey G Shaffer, Daouda Sanogo, Soumba Keita, Moussa Keita, Fousseyni Kane, Bourama Traore, Djeneba Dabitao, Aissata Kone, Cheick Oumar Doumbia, Joseph Keating, Joshua Yukich, Helle H Hansson, Alyssa E Barry, Mahamadou Diakité, Michael Alifrangis, Seydou Doumbia. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.01.2024. 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.