JMIR Research Protocols

Impact Factor (2023): 1.4
Volume 9 (2020), Issue 4 ISSN 1929-0748 Editor in Chief: Xiaomeng (Simone) Ma, PhDc, MS, BS

Contents

Protocols

Original Paper

Protocol

Gamified Mobile Computerized Cognitive Behavioral Therapy for Japanese University Students With Depressive Symptoms: Protocol for a Randomized Controlled Trial

Kengo Yokomitsu[1], PhD; Tomonari Irie[2], MA; Mayu Sekiguchi[3], PhD; Ayako Shimizu[4], MA; Hirofumi Matsuoka[5], PhD; Sally Nicola Merry[6], MBChB, MD, FRANZCP, CCAP; Karolina Stasiak[6], PhD

[1]College of Comprehensive Psychology, Ritsumeikan University, Osaka, Japan

[2]School of Education and Culture, Hokusho University, Hokkaido, Japan

[3]School of Psychological Science, Health Sciences University of Hokkaido, Hokkaido, Japan

[4]HIKARI Lab, Inc, Tokyo, Japan

[5]School of Dentistry, Health Sciences University of Hokkaido, Hokkaido, Japan

[6]Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

Corresponding Author:
Kengo Yokomitsu, PhD
College of Comprehensive Psychology
Ritsumeikan University
2-150 Iwakura-cho, Ibaraki
Osaka, 567-8570
Japan
Phone: 81 72 665 2490
Email: k-yoko@fc.ritsumei.ac.jp

Abstract

Background: Evidence shows that computerized self-help interventions are effective for reducing symptoms of depression. One such intervention, SPARX, is a gamified mobile computerized cognitive behavioral therapy (cCBT) developed for adolescents in New Zealand, which was shown to be as effective as usual care for young people with mild-to-moderate symptoms of depression. However, gamified cCBT has not yet been tested in Japan.

Objective: This trial is designed to investigate whether a Japanese-adapted version of SPARX improves depressive symptoms in Japanese university students with mild-to-moderate depressive symptoms.

Methods: In this 7-week, multicenter, stratified, parallel-group, superiority randomized trial, participants will be allocated to either a treatment condition (SPARX) or a wait-list control condition. SPARX is a fully automated program, which will be delivered to the mobile phone or tablet device of the participants. SPARX is designed as an interactive fantasy game to guide the user through seven modules that teach key CBT strategies. All participants will be recruited from universities via advertisements on online bulletin boards, the campus newspaper, and posters. Participants in the treatment condition will use the SPARX program weekly. The primary outcome is the reduction of depressive symptoms (using Patient Health Questionnaires-9) measured at baseline and weekly: once after the 7-week intervention and once at a 1-month follow-up. Secondary outcomes include satisfaction with the program and satisfaction with life, measured by the Satisfaction With Life Scale; positive and negative moods, measured by the Profile of Mood States Second Edition; social functioning, measured by the EuroQol Instrument; rumination, measured by the Ruminative Responses Scale; and coping, measured by the Brief Coping Orientation to Problem Experienced Inventory.

Results: This study received funding from The Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, and obtained institutional review board approval in September 2019. Data collection began in April 2019.

Conclusions: Results of this trial may provide further evidence for the efficacy of gamified cCBT for the treatment of depression and, specifically, provide support for using SPARX with Japanese university students.

Trial Registration: Japan Primary Registries Network UMIN000034354; https://tinyurl.com/uu7xd77

International Registered Report Identifier (IRRID): DERR1-10.2196/15164

(JMIR Res Protoc 2020;9(4):e15164) doi:10.2196/15164

KEYWORDS

SPARX; Japan; university students; depressive symptoms

Introduction

Background

Depression is a mental health problem commonly experienced by university students. A systematic review examining rates of depression among university students from a range of countries, including North American, European Union, East Asian, and Middle Eastern countries, revealed the weighted mean prevalence rate of depressive disorders of 30.6% [1], which is considerably higher than that reported for the general population (6.6%) [2]. Depression and other mental health concerns such as anxiety and stress negatively impact academic performance [3-5], which may result in students dropping out of university [6] and potentially affect their future career prospects. For example, in Japan, only 1.7% of individuals that dropped out of university reported eventually working as full-time permanent employees for three or more years, whereas 60% of university and graduate school graduates obtained regular employment [7]. Therefore, the mental well-being of tertiary students is not only a health care issue but also an issue that influences the overall socioeconomic welfare of a society.

In Japan, 126 of the 135 universities (93%) have some established support systems to address the mental health needs of university students [8]. However, reaching out to professional mental health services among students is relatively uncommon [9,10] owing to several factors, including skepticism about treatment effectiveness, lack of awareness of help options, and lack of perceived need [11,12]. Therefore, access to university mental health services should be adjusted so that it is both easier and more acceptable [13]. Outside Japan, there have been many trials of internet-based and computer-delivered interventions in recent years, demonstrating significant improvements in depression compared to an inactive control (pooled standardized mean difference –0.43, 95% CI –0.63 to –0.22) [14]. Digital interventions offer important advantages: they are accessible from various locations, they can be a form of outreach to individuals who might not otherwise access traditional face-to-face services [15,16] and are more cost-effective than face-to-face treatment [17]. Moreover, digital interventions can facilitate learning and retention because users can return to the program at their convenience to practice the content more than once [18]. Given that games are familiar and particularly popular with young people in Japan [19,20], game-based interventions have the potential to be engaging and valuable self-help tools to support university students' mental health [21,22].

In earlier studies, a variety of game-based interventions for mental health suggested potential benefits for mental and physical symptoms [23-25]. At present, the Smart, Positive, Active, Realistic, X-factor thoughts (SPARX) program [26] is the only available game-based intervention developed to specifically deliver cognitive behavioral therapy (CBT) for adolescent users with depressive symptoms [22,27]. SPARX was created, evaluated, and implemented nationally in New Zealand for use by adolescents with mild-to-moderate depressive symptoms [26]. Furthermore, SPARX was used as a depression prevention intervention in a large Australian school trial to reduce depressive symptoms in adolescents prior to final year examinations [28]. In Japan, a smartphone-based CBT program (Kokoro-app) was demonstrated to be effective for depression in a clinical sample [29]; however, there have been no trials of game-based CBT for depression in Japanese university students conducted to date. It is anticipated that SPARX may be well received by university students experiencing depressive symptoms who do not have contact with professional services. However, a formal randomized controlled trial (RCT) is warranted to establish whether SPARX is effective in this population.

Objectives

The primary objective of this trial is to examine whether SPARX improves depressive symptoms in Japanese university students who present mild-to-moderate symptoms. We hypothesize that SPARX will be more effective for improving depressive symptoms than a wait-list condition. In the systematic review of trials of internet-based and computer-delivered interventions for university students [14], although half of the included trials used an inactive control group (wait-list or no treatment), the other half used an active control (eg, attention training or providing educational information on anxiety and depressive disorders). This will be the first clinical trial of SPARX for university students in Japan. Therefore, we will examine the difference between a SPARX group and a wait-list control group.

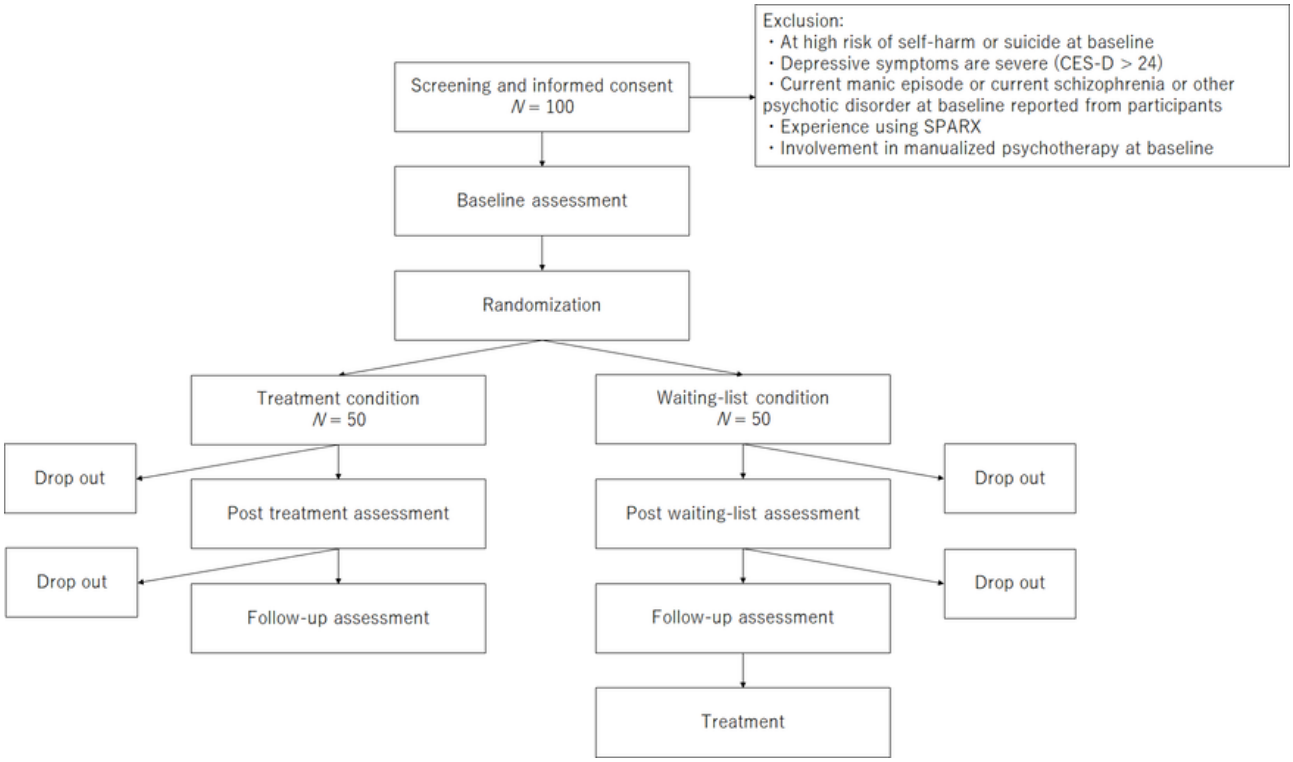
The secondary objectives are to determine whether SPARX delivered to Japanese students with mild-to-moderate depressive symptoms leads to (1) enhanced positive mood and reduced negative mood, (2) improved satisfaction with life, (3) improved health-related quality of life, (4) reduced depressive rumination, and (5) improved coping skills.

Methods

Study Design

This is a multicenter trial and follows a stratified randomized, parallel-group superiority design. Participants will be randomized 1:1 to an intervention arm, which will consist of the game-based computerized CBT program SPARX or a wait-list control condition. Postintervention assessments will be carried out 7 weeks after baseline assessment. The participant flow through the trial is shown in Figure 1.

Figure 1. Participant flow of this trial. CES-D: Center for Epidemiological Studies Depression; SPARX: Smart, Positive, Active, Realistic, X-factor thoughts.



Setting

This multicenter study will be conducted at three universities: Health Sciences University of Hokkaido, Hokusho University, and Ritsumeikan University. We plan to recruit participants from April 2019 through March 2020.

Participants

A total of 100 participants will be sought to take part in this trial. Participants will be included in the trial if they meet the following inclusion criteria at the face-to-face or telephone eligibility interview: (1) mild-to-moderate depressive symptoms on the Center for Epidemiological Studies Depression Scale (CES-D) [30] with scores between 16 and 24 [31,32], (2) age of 18-30 years, (3) attending one of the participating universities, and (4) owning and being able to use a smartphone/tablet device. These inclusion criteria differ from those of the first SPARX study [26]. Namely, we include university students aged 18-30 years instead of high school students, who were recruited for the first SPARX study [26], because when the Japanese version of SPARX was first developed, this app was designed with adult users in mind, including university students, rather than high school students. In addition, we decided to conduct the current study based on the contemporary reality of Japanese university students, because a study conducted by the Ministry of Education, Culture, Sports, Science and Technology of Japan in 2018 [33] found that Japanese university students aged 18-30 years accounted for approximately 99.9% of the total university students in Japan.

Participants will be excluded if they have: (1) high risk of self-harm or suicide at study entry based on a self-report; (2) severe depressive symptoms, as indicated by a CES-D score>24 [31,32]; (3) current manic episodes or current schizophrenia or other psychotic disorders at baseline as reported by a participant; (4) previous experience of using SPARX; and (5) involvement in manualized psychotherapy at baseline. These exclusion criteria are the same as those adopted in the first SPARX study [26] as reference. Merry et al [26] suggested that people exhibiting severe depressive symptoms or at high risk of self-harm or suicide should be excluded because they could not provide a self-help resource as a viable option. In addition, individuals experiencing current manic episodes, schizophrenia, or other psychotic disorders, or are involved in manualized psychotherapy at baseline were excluded because they require medical and psychiatric treatment suitable for their symptoms rather than SPARX or are already receiving suitable psychiatric treatment. If we do not reach the sample size required to test the primary hypothesis of this study, we will extend recruitment to other interested universities and will consider using online recruitment.

Clinical Safety Procedures

For participants indicating a high risk of self-harm or suicide at screening and at any time during this trial, the researchers will provide information on available mental health care services that the participants can access and will ask the participant to cease participation in this trial. However, participants will be provided with ongoing access to SPARX even if they decide to withdraw from the trial. In relation to ancillary care, the participants will be allowed to contact the researchers at any point during the trial. If any event requiring additional care occurs, a researcher will inform the individual at the contact address or telephone number of the appropriate medical or psychological care and mental health service as provided by the participant's health insurance.

Study Procedure

All participants will be recruited from the abovementioned universities via advertisements on online bulletin boards, campus newspapers, and posters. Interested students who contact the researchers via email will be provided with further information related to the study and will be given time to consider their participation. If interested, they will be screened using Google forms (developed by the authors for this study) for eligibility using the Japanese version of the CES-D [30]. The cutoff score of CES-D was verified in Japanese samples [30]. If the score is 16-24, the researcher will continue with the eligibility interview to confirm that all other criteria are met. At the end of the interview, written consent will be obtained. Those who score less than 15 or above 25 will be excluded from the trial but will be provided with free access to SPARX, according to the needs of the participant. Those whose score is 25 or more, indicating severe depressive symptoms [31,32], will be provided with information about available mental health care services.

For participants in the treatment condition, SPARX will be delivered on their mobile phone or tablet device. They will be asked to play the SPARX program weekly at a designated university location (eg, a vacant laboratory or conference room) with one of the researchers (or a research assistant) present in another room (eg, a neighboring laboratory) for the purpose of checking SPARX use and being available to answer any questions from the participant. Each weekly session with SPARX is expected to take about 30 min.

Participants in the wait-list condition group will receive access to SPARX after a 1-month follow-up assessment of the treatment condition group. They will be asked to fill out assessments at the same time as the assessments of the treatment condition group at experimental laboratories, in conference rooms at each university, or on online Google forms. They will not be granted access to SPARX until all assessments have been completed.

To address potential issues of contamination, we forbid participants from talking about any theme of SPARX, such as the use of SPARX and any impressions of this study, during trial participation because other students in their university campus are participating in the current trial.

Ethics

This study protocol was reviewed and approved by the institutional review board of Hokusho University (2017-021; February 20, 2018), Ritsumeikan University (2018-028; September 19, 2018), and Health Sciences University of Hokkaido (18-006, December 8, 2018). Furthermore, this study has been registered with UMIN (000034354). Details of this trial will be explained by the researchers to the participants before study participation. Written informed consent will be obtained from participants prior to entering them into the trial. All outcome data, including written informed consent, will be securely stored at the participating university (eg, Ritsumeikan University) in a locked cabinet. All identifying information will be removed, and the participants will only be identified by a unique participant number.

Randomization

Participants (N=100) who fulfil all eligibility criteria will be randomized at a 1:1 ratio to either the SPARX intervention group or the wait-list control condition. The randomization sequence will be generated by the trial statistician and stratified by site. The final randomization lists will be computer-generated and concealed in a secure study database until the end of data collection/data lock.

Sample Size

A meta-analysis of RCTs examining the effects of CBT in comparison to wait-list on depression in university students showed an effect size of 1.13 [34]. However, the effect sizes were smaller (0.56 [35] and 0.64 [36]) for studies in adults compared to those of internet-based and other computerized psychological treatments with wait lists. A meta-analysis of RCTs to examine the effectiveness of game-based digital interventions showed an overall effect size of 0.66 (against a wait-list control condition), 0.47 (with a target population of adolescent users), and 0.41 (type of game: psychoeducation and training, including the SPARX study) [22]. Additionally, a meta-analysis of RCTs to examine the effectiveness of technology-delivered interventions for depression and anxiety in children and adolescents showed an effect size of 0.45 compared to the wait-list group [37]. Considering these findings, we assumed a similar effect size (0.60) to that of internet-based and other computerized psychological treatments or game-based digital interventions against wait-list [22,35,36] for the reduction of primary outcomes assessed using Patient Health Questionnaire-9 (PHQ-9).

A power analysis showed that a target sample size of 90 was needed to detect an effect size of 0.60 with alpha=.05 and a power of 0.80 for a two-tailed test. Considering an anticipated 10% dropout [38], we will aim to recruit at least 100 (50 per group) participants to test the primary hypothesis of this study.

Intervention

The treatment condition in this study consists of a game-based computerized CBT program called SPARX. It takes the form of an interactive fantasy game designed to take the user through seven modules that teach key CBT strategies. When the user begins SPARX, they meet the character of the Guide who introduces them to the “game world” and the challenges ahead, after which the user selects their avatar and begins the quest. Each module finishes with the avatar returning to the Guide who then explains how the “game skills” can be used in “real life”. The first module includes brief psychoeducation related to depression, with a brief introduction of relaxation (slow and controlled breathing) and hope. The remaining modules cover the following: activity scheduling/behavioral activation, interpersonal skills (communication, assertiveness, and negotiation), problem solving, cognitive restructuring (identifying common cognitive distortions and challenging negative thoughts), mindfulness, and relapse prevention (Textbox 1). Each module is designed to take about 20-30 min and the user is encouraged to practice the skills in the interim (the equivalent of homework).

Textbox 1. Outline of content and core skills covered in each level of Smart, Positive, Active, Realistic, X-factor thoughts (SPARX).

- Level 1 (finding hope): psychoeducation about depression, introduction to the cognitive behavioral therapy model, introducing gloomy negative automatic thoughts and "hope" (people recover from depression), relaxation (controlled breathing)
- Level 2 (being active): activity scheduling, behavioral activation, relaxation (progressive muscle relaxation), basic communication, and interpersonal skills
- Level 3 (dealing with emotions): dealing with anger and hurt feelings, interpersonal skills (assertiveness, listening, and negotiation)
- Level 4 (overcoming problems): problem solving, cognitive restructuring (identifying smart, positive, active, realistic, X-factor thoughts)
- Level 5 (recognizing unhelpful thoughts): cognitive restructuring (recognizing different types of gloomy negative automatic thoughts)
- Level 6 (challenging unhelpful thoughts): cognitive restructuring (learning to challenge or swap negative thoughts for helpful ones), interpersonal skills continued (negotiation skills)
- Level 7 (bringing it all together): recap of all skills, mindfulness (tolerating distress), relapse prevention

The Japanese version of SPARX was developed by the Japanese companies HIKARI Lab, Inc, and SmileBoom Co, Ltd, with approval of the developers of the original version. The original English language was translated into Japanese by the fourth author (AS) who has a master's degree in clinical psychology and is a native Japanese speaker and a fluent speaker of English. A key difference between the original and Japanese versions was the change of the Guide character's gender from male to female. In the original version, the Guide was male, stylized to look like a Maori (the indigenous people of New Zealand) chief to give a powerful impression. However, the developers in the Japanese version were concerned that this might not suit Japanese audiences and instead adopted a female character wearing a white national costume, with a more gentle and maternal expression (Figure 2). Other modifications include the following: for all SPARX characters, the face and eyes were slightly rounded and the color schemes were made brighter to be more reminiscent of the Japanese anime esthetic; login and menu screens were slightly changed to facilitate use; by using a notification function, the user can obtain information related to mental health once a week; how to process SPARX is displayed; and as the user plays the game and proceeds through the stages, the sky color becomes brighter. There were also some necessary content changes to make the program more applicable to Japanese adult and university students as users (eg, original version: assignment/homework anxiety; Japanese version: work presentation anxiety). With the exception of these minor and complementary fixes in developing the Japanese version of SPARX, no major changes to the original were made, such as a change of the main story of SPARX. The original version was delivered on a user's personal computer, whereas the Japanese version is delivered on a mobile phone or tablet device.

Figure 2. Guide characters in the SPARX original (left) and Japanese (right) versions.

Wait-List Control

Participants randomized to the control group will be placed on a wait list for 11 weeks; that is, until after the 1-month follow-up assessment of the group assigned to the intervention. The participants on the wait list will then be offered the opportunity to take part in the SPARX program in the same way as those in the intervention group.

Measures

The primary outcome (Japanese version of PHQ-9) will be measured at baseline (T1), weekly throughout the interventions (T2-T7), at postassessment 1 week after the 7 intervention weeks (T8), and at 1-month follow-up (T9). Other assessments will be conducted at preassessment (T1), postassessment (T8), and 1-month follow-up (T9). Table 1 shows the schedule of assessments at various time points. At baseline (T1), demographics data will be collected, including gender, age, major in university, year of study, residence status; presence or absence of visits to psychiatric and counseling services; and presence or absence of the use of a gamified psychological program.

Table 1. Overview of measurements.

Outcome	Scale	T1[a]	T2-T7[b]	T8[c]	T9[d]
Demographics	Developed by the authors	✓			
Depressive symptoms	PHQ-9[e]	✓	✓	✓	✓
Satisfaction with life	SWLS[f]	✓		✓	✓
Negative and positive moods	POMS-2[g]	✓		✓	✓
Social functions	EQ-5D[h]	✓		✓	✓
Rumination	RSS[i]	✓		✓	✓
Coping	Brief COPE[j]	✓		✓	✓
Stressor	DLSS[k]	✓		✓	✓
Satisfaction and acceptability	Developed by the authors				✓

[a]T1: baseline assessment (preassessment).

[b]T2-T7: assessment at each level.

[c]T8: assessment 1 week after the program.

[d]T9: assessment at 1-month follow-up.

[e]PHQ-9: Japanese version of Patient Health Questionnaire-9.

[f]SWLS: Satisfaction With Life Scale.

[g]POMS-2: Profile of Mood Scale, second edition.

[h]EQ-5D: Japanese EuroQol instrument.

[i]RSS: Japanese version of the Ruminative Responses Scale.

[j]Brief COPE: Brief Coping Orientation to Problem Experienced inventory.

[k]DLSS: Daily Life Stress Scale for undergraduates.

Primary Outcome Measure

The primary outcome measure is the Japanese version of PHQ-9. The PHQ-9 is a self-reported questionnaire for assessing depressive symptoms of the preceding 2 weeks [39]. The questionnaire consists of nine items evaluated using a 4-point Likert scale to indicate the extent to which the participant agrees with the values expressed in each item (0, not at all; 3, nearly every day). Higher scores indicate more severe depressive symptoms. In an earlier study [39], this scale demonstrated good internal consistency (alpha=.93) and good convergent validity (correlation coefficient with the Kessler Psychological Distress Scale: r=.81) in a Japanese clinical population.

Secondary Outcome Measures

The following sections will discuss the secondary measures used.

Japanese Version of the Profile of Mood States Second Edition (POMS-2)

This is a self-reported questionnaire for assessing positive and negative moods during the preceding week [40]. The questionnaire consists of 35 items evaluated on a 5-point Likert scale (0-4). For this study, we will use total mood disturbance; higher scores indicate more negative mood states. The reliability and validity of POMS-2 for a Japanese population have been reported [40].

The Satisfaction With Life Scale (SWLS)

This is a self-reported questionnaire for assessing satisfaction with life [41]. The questionnaire consists of five items evaluated on a 7-point Likert-type scale (1-7). Higher scores indicate more satisfaction with life. Its reliability and validity for a Japanese population have been reported [41].

The Japanese EuroQol Instrument (EuroQol)

This is a self-reported questionnaire for assessing health-related quality of life [42]. The questionnaire consists of five items evaluated on a 5-point Likert-type scale (1-5). Higher scores indicate worse health-related quality of life. Its reliability and validity for a Japanese population have been reported [43].

The Japanese Version of the Ruminative Responses Scale (RSS)

This is a self-reported questionnaire for assessing the frequency of depressive rumination [44]. The questionnaire consists of 22 items evaluated on a 7-point Likert-type scale (1-7). Higher scores indicate more frequent depressive rumination. Its reliability and validity for a Japanese population have been reported [45].

The Brief Coping Orientation to Problem Experienced Inventory (Brief COPE)

This is a self-rated questionnaire for assessing coping skills [46]. The questionnaire consists of 28 items evaluated on a 4-point Likert-type scale (1-4). Higher scores represent high levels of coping skills. Its reliability and validity for a Japanese population have been reported [47].

The Daily Life Stress Scale for Undergraduates (DLSS)

This is a self-reported questionnaire for assessing daily life stressors for undergraduates [48]. The questionnaire consists of 23 items evaluated on a 5-point Likert-type scale (1-5). Higher scores indicate more frequent stressors. Its reliability and validity for a Japanese population have been reported [48].

The final secondary outcome will be a measure of the satisfaction and acceptability of the intervention. All participants will be asked to rate four statements related to the perceived helpfulness, satisfaction with use, and the depth of understanding of the content of SPARX on an 11-point scale (0-10). Participants will also be able to write free text comments. The following questions will be asked: "How useful was this program for you [0 (not at all), 10 (very useful)]? Specifically, which components of this program were useful for you (write free text comments)?" "How satisfied are you with this program [0 (not at all), 10 (very satisfied)]?" "Please tell me about the pros and cons of this program (write free text comments)?" "How fun was this program for you [0 (not at all), 10 (very fun)]?" These items were developed specifically for the current trial.

Statistical Analyses

The analysis for primary outcome measures will be conducted using a linear mixed model (LMM). The LMM approach was selected because of its strength in accommodating missing data and its ability to incorporate random effects into analyses. In these analyses, the dependent variable is the PHQ-9 score, and the independent variables are assignment (treatment condition, wait-list condition) and time [preassessment (T1), each session assessment (T2-T7), post-assessment (T8), and follow-up assessment (T9)], with interaction of the assignment and time as fixed-effect variables and participants as a random-effects variable. Secondary outcomes (outcome measures, process measures, and evaluation of the program) will be analyzed in the same way as the primary outcome. Subgroup analyses are not planned. For all analyses, $P<.05$ will be inferred as statistically significant. Up-to-date versions of R software (R Foundation for Statistical Computing, Vienna, Austria) will be used to analyze the quantitative data.

In the satisfaction/acceptability measure, free comments will be analyzed using thematic analysis.

Results

The primary outcome is the presence of depressive symptoms at 1-month follow-up. Several secondary outcomes will be measured, such as positive and negative mood, satisfaction with life, health-related quality of life, depressive rumination, and coping skills. In addition, satisfaction and acceptability of the intervention will be measured, such as the perceived helpfulness, satisfaction with use, and the depth of understanding of the content of SPARX. Dropout rates will be measured to study the feasibility of SPARX.

This study received funding from The Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, and obtained institutional review board approval in September 2019. Data collection began in April 2019. Study enrollment is still ongoing. As of April 2019, 60 subjects have been recruited and baseline data were collected. Moreover, 24 subjects have already been allocated to the treatment group and have begun the intervention. We expect to complete the study by March 2020, followed by data analysis and submission of the final report by the end of 2020.

Irrespective of the results, the results of this trial will be published in a peer-reviewed journal and communicated through presentations at national and international conferences. Study participants will be informed about the trial results via a plain-language summary of the results that will be sent to them. Academic papers and summary reports will be provided to the funders.

Discussion

SPARX is a game-based intervention that was developed to deliver CBT for adolescents with depressive symptoms [22,27]. This program has been demonstrated to be effective for high school students with mild-to-moderate symptoms of depression in New Zealand [26]. The aim of this trial is to investigate the efficacy of the program in Japanese university students using a superiority RCT comparing SPARX with a wait-list control. The findings of this trial are expected to add to evidence of the efficacy of the game-based computerized CBT for young people with depressive symptoms. These results will provide important knowledge for implementation of computerized CBT into existing health support services for university students in the future. However, in this trial, for the purpose of checking SPARX use and being able to answer any questions from a participant, participants in the treatment condition play the SPARX program at a designated university location with one of the researchers present in another room. Therefore, we are not examining the pure effect of SPARX as a self-help tool. This procedure could influence the intervention effect or dropout ratio. In addition, to clarify the benefits and limitations of the effect of SPARX on the university students' depressive symptoms, the effects of SPARX in comparison to an active control will need to be assessed in the future.

Authors' Contributions

All authors contributed equally to the study design by providing relevant intellectual input on all aspects of the study within several detailed discussions and meetings. MS and HM raised the research funding. AS developed the Japanese version of SPARX. MS and HM developed the original version of SPARX. KY drafted the manuscript. All authors revised the manuscript. All authors approved the version to be published.

Conflicts of Interest

AS developed the Japanese version of SPARX and would benefit financially from its commercialization. SM and KS are the codevelopers of the original version of SPARX in New Zealand. SM and KS receive commercialization royalties from any licensed use of SPARX, including Japan.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 293 KB - resprot_v9i4e15164_app1.pdf]

References

1. Ibrahim AK, Kelly SJ, Adams CE, Glazebrook C. A systematic review of studies of depression prevalence in university students. J Psychiatr Res 2013 Mar;47(3):391-400. [doi: 10.1016/j.jpsychires.2012.11.015] [Medline: 23260171]
2. Gonalez O, Berry J, Mcknighty-Eily I, Strine T, Edwards V, Lu H, et al. Current depression among adults: United States, 2006 and 2008. MMWR Morb Mortal Wkly Rep 2010;59(38):1229-1235. [doi: 10.1037/e664792010-002] [Medline: 20881934]
3. Richardson M, Abraham C, Bond R. Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychol Bull 2012;138(2):353-387. [doi: 10.1037/a0026838] [Medline: 22352812]
4. Riglin L, Petrides K, Frederickson N, Rice F. The relationship between emotional problems and subsequent school attainment: A meta-analysis. J Adolesc 2014 Jun;37(4):335-346. [doi: 10.1016/j.adolescence.2014.02.010] [Medline: 24793380]
5. Steptoe A, Tsuda A, Tanaka Y, Wardle J. Depressive symptoms, socio-economic background, sense of control, and cultural factors in university students from 23 countries. Int J Behav Med 2007 Jun;14(2):97-107. [doi: 10.1007/bf03004175] [Medline: 17926438]
6. Ministry of Education, Culture, Sports, Science, and Technology, Japan. MEXT, Japan. 2016. About the current situation of students withdrawing from abroad or taking leave of absence URL: http://warp.ndl.go.jp/info:ndljp/pid/11293659/www.mext.go.jp/b_menu/houdou/26/10/1352425.htm [accessed 2019-07-22]
7. The Japan Institute for Labour PolicyTraining. JILPT survey series. 2015. Research on employment and consciousness of dropouts of university URL: https://www.jil.go.jp/english/reports/documents/jilpt-research/no.138.pdf [accessed 2019-07-22]
8. Japan Student Services Organization. Japan Student Services Organization. 2017. Survey of the status of student support efforts at universities in Japan URL: https://www.jasso.go.jp/about/statistics/torikumi_chosa/__icsFiles/afieldfile/2017/04/14/h27torikumi_houkoku.pdf [accessed 2019-07-22]
9. Goodwin J, Behan L, Kelly P, McCarthy K, Horgan A. Help-seeking behaviors and mental well-being of first year undergraduate university students. Psychiatry Res 2016 Dec 30;246:129-135. [doi: 10.1016/j.psychres.2016.09.015] [Medline: 27693865]
10. Hunt J, Eisenberg D. Mental health problems and help-seeking behavior among college students. J Adolesc Health 2010 Jan;46(1):3-10. [doi: 10.1016/j.jadohealth.2009.08.008] [Medline: 20123251]
11. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. Med Care 2007 Jul;45(7):594-601. [doi: 10.1097/MLR.0b013e31803bb4c1] [Medline: 17571007]
12. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
13. Mowbray CT, Megivern D, Mandiberg JM, Strauss S, Stein CH, Collins K, et al. Campus mental health services: recommendations for change. Am J Orthopsychiatry 2006 Apr;76(2):226-237. [doi: 10.1037/0002-9432.76.2.226] [Medline: 16719642]
14. Davies EB, Morriss R, Glazebrook C. Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: a systematic review and meta-analysis. J Med Internet Res 2014 May 16;16(5):e130 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]
15. Taylor CB, Luce KH. Computer- and Internet-Based Psychotherapy Interventions. Curr Dir Psychol Sci 2016 Jun 24;12(1):18-22. [doi: 10.1111/1467-8721.01214]
16. Fleming T, Merry S. Youth Work Service Providers' Attitudes Towards Computerized CBT for Adolescents. Behav Cogn Psychother 2012 May 17;41(3):265-279. [doi: 10.1017/S1352465812000306] [Medline: 22591811]
17. Griffiths KM, Christensen H. Internet-based mental health programs: a powerful tool in the rural medical kit. Aust J Rural Health 2007 Apr;15(2):81-87. [doi: 10.1111/j.1440-1584.2007.00859.x] [Medline: 17441815]
18. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
19. Tsuchiya M, Momma H, Sekiguchi T, Kuroki K, Kanazawa K, Watanabe M, et al. Excessive Game Playing Is Associated with Poor Toothbrushing Behavior among Athletic Children: A Cross-Sectional Study in Miyagi, Japan. Tohoku J Exp Med 2017 Feb;241(2):131-138 [FREE Full text] [doi: 10.1620/tjem.241.131] [Medline: 28190825]

20. Mak K, Lai C, Watanabe H, Kim D, Bahar N, Ramos M, et al. Epidemiology of internet behaviors and addiction among adolescents in six Asian countries. Cyberpsychol Behav Soc Netw 2014 Nov;17(11):720-728. [doi: 10.1089/cyber.2014.0139] [Medline: 25405785]
21. Kauer SD, Mangan C, Sanci L. Do Online Mental Health Services Improve Help-Seeking for Young People? A Systematic Review. J Med Internet Res 2014 Mar 04;16(3):e66 [FREE Full text] [doi: 10.2196/JMIR.3103] [Medline: 24594922]
22. Li J, Theng Y, Foo S. Game-based digital interventions for depression therapy: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2014 Aug;17(8):519-527 [FREE Full text] [doi: 10.1089/cyber.2013.0481] [Medline: 24810933]
23. Khazaal Y, Chatton A, Dieben K, Huguelet P, Boucherie M, Monney G, et al. Reducing Delusional Conviction through a Cognitive-Based Group Training Game: A Multicentre Randomized Controlled Trial. Front Psychiatry 2015 Apr;6:66 [FREE Full text] [doi: 10.3389/fpsyt.2015.00066] [Medline: 25972817]
24. Leutwyler H, Hubbard E, Cooper B, Dowling G. The Impact of a Videogame-Based Pilot Physical Activity Program in Older Adults with Schizophrenia on Subjectively and Objectively Measured Physical Activity. Front Psychiatry 2015 Dec 21;6:180. [doi: 10.3389/fpsyt.2015.00180] [Medline: 26733891]
25. Tárrega S, Castro-Carreras L, Fernández-Aranda F, Granero R, Giner-Bartolomé C, Aymamí N, et al. A Serious Videogame as an Additional Therapy Tool for Training Emotional Regulation and Impulsivity Control in Severe Gambling Disorder. Front Psychol 2015;6:1721 [FREE Full text] [doi: 10.3389/fpsyg.2015.01721] [Medline: 26617550]
26. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012 Apr 19;344(3):e2598. [doi: 10.1136/bmj.e2598] [Medline: 22517917]
27. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious Games and Gamification for Mental Health: Current Status and Promising Directions. Front Psychiatry 2016 Jan;7:215 [FREE Full text] [doi: 10.3389/fpsyt.2016.00215] [Medline: 28119636]
28. Perry Y, Werner-Seidler A, Calear A, Mackinnon A, King C, Scott J, et al. Preventing Depression in Final Year Secondary Students: School-Based Randomized Controlled Trial. J Med Internet Res 2017 Nov 02;19(11):e369 [FREE Full text] [doi: 10.2196/jmir.8241] [Medline: 29097357]
29. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone Cognitive Behavioral Therapy as an Adjunct to Pharmacotherapy for Refractory Depression: Randomized Controlled Trial. J Med Internet Res 2017 Nov 03;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602] [Medline: 29101095]
30. Shima S, Shikano T, Kitamura T, Asai M. New self-rating scales for depression [in Japanese]. Japanese Journal of Psychiatry 1985;27(6):717-723.
31. Schulberg HC. Assessing Depression in Primary Medical and Psychiatric Practices. Arch Gen Psychiatry 1985 Dec 01;42(12):1164. [doi: 10.1001/archpsyc.1985.01790350038008] [Medline: 4074109]
32. Cho MJ, Nam JJ, Suh GH. Prevalence of symptoms of depression in a nationwide sample of Korean adults. Psychiatry Res 1998 Dec;81(3):341-352. [doi: 10.1016/s0165-1781(98)00122-x] [Medline: 9925185]
33. Ministry of Education, Culture, Sports, Science, and Technology, Japan. MEXT Japan. 2018. The Study Basic Research: The summary of results in 2018 URL: http://www.mext.go.jp/component/b_menu/other/__icsFiles/afieldfile/2018/12/25/1407449_3.pdf [accessed 2019-07-22]
34. Cuijpers P, Cristea IA, Ebert DD, Koot HM, Auerbach RP, Bruffaerts R, et al. Psychological treatment of depression in college students: a metaanalysis. Depress Anxiety 2015 Dec 18;33(5):400-414. [doi: 10.1002/da.22461] [Medline: 26682536]
35. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009 Dec;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
36. So M, Yamaguchi S, Hashimoto S, Sado M, Furukawa TA, McCrone P. Is computerised CBT really helpful for adult depression?-A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity. BMC Psychiatry 2013 Apr 15;13(1):113. [doi: 10.1186/1471-244X-13-113] [Medline: 23587347]
37. Grist R, Croker A, Denne M, Stallard P. Technology Delivered Interventions for Depression and Anxiety in Children and Adolescents: A Systematic Review and Meta-analysis. Clin Child Fam Psychol Rev 2018 Sep 18;22:147-171. [doi: 10.1007/s10567-018-0271-8] [Medline: 30229343]
38. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
39. Doi S, Ito M, Takebayashi Y, Muramatsu K, Horikoshi M. Factorial validity and invariance of the Patient Health Questionnaire (PHQ)-9 among clinical and non-clinical populations. PLoS ONE 2018 Jul 19;13(7):e0199235. [doi: 10.1371/journal.pone.0199235] [Medline: 30024876]
40. Yokoyama K. Profile of Mood States Second Edition Japanese. Tokyo: Kanekoshobo; 2015.
41. Sumino Z. The development of the Japanese version of the Satisfaction With Life Scale. 1994 Presented at: Annual Convention of the Japanese Association of Educational Psychology; September 28, 1994; Kyoto p. 192. [doi: 10.20587/pamjaep.36.0_192]
42. Nishimura S, Tsuchiya A, Hisashige A, Ikegami N, Ikeda S. The development of the Japanese EuroQol Instrument. Iryo To Shakai 1998;8(1):109-123. [doi: 10.4091/iken1991.8.1_109]

43. Tsuchiya A, Hasegawa T, Nishimura S, Hisashige A, Ikegami N, Ikeda S. A Validity Study of the Japanese EuroQol Instrument. Iryo To Shakai 1998;8(1):67-77. [doi: 10.4091/iken1991.8.1_67]
44. Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: A psychometric analysis. Cogn Ther Res 2003;27(3):247-259. [doi: 10.1023/A:1023910315561]
45. Hasegawa A. Translation and Initial Validation of the Japanese Version of the Ruminative Responses Scale. Psychol Rep 2013 Jun;112(3):716-726. [doi: 10.2466/02.08.PR0.112.3.716-726] [Medline: 2425067]
46. Carver CS. You want to measure coping but your protocol's too long: consider the brief COPE. Int J Behav Med 1997;4(1):92-100. [doi: 10.1207/s15327558ijbm0401_6] [Medline: 16250744]
47. Otsuka Y, Sasaki T, Iwasaki K, Mori I. Working hours, coping skills, and psychological health in Japanese daytime workers. Ind Health 2009 Jan;47(1):22-32 [FREE Full text] [doi: 10.2486/indhealth.47.22] [Medline: 19218754]
48. Shima N. Examination of the Daily Life Stress Scale for undergraduates. Japanese Chukyo University Department of Sociology Bulletin 1999;14(1):69-83.

Abbreviations

Brief COPE: Brief Coping Orientation to Problem Experienced inventory
CBT: cognitive behavioral therapy
CES-D: Center for Epidemiological Studies Depression
DLSS: Daily Life Stress Scale for undergraduates
EuroQol: EuroQol Instrument
LMM: linear mixed model
PHQ-9: Patient Health Questionnaire-9
POMS-2: Profile of Mood States Second Edition
RSS: Ruminative Responses Scale
SPARX: Smart, Positive, Active, Realistic, X-factor thoughts
SWLS: Satisfaction With Life Scale
UMIN: University Hospital Medical Information Network

Edited by C Hoving; submitted 05.08.19; peer-reviewed by K Lawrence, M Lucassen; comments to author 27.09.19; revised version received 12.12.19; accepted 18.12.19; published 07.04.20.

Please cite as:
Yokomitsu K, Irie T, Sekiguchi M, Shimizu A, Matsuoka H, Merry SN, Stasiak K
Gamified Mobile Computerized Cognitive Behavioral Therapy for Japanese University Students With Depressive Symptoms: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e15164
URL: https://www.researchprotocols.org/2020/4/e15164
doi:10.2196/15164
PMID:32254045

Protocol

Road to Hierarchical Diabetes Management at Primary Care (ROADMAP) Study in China: Protocol for the Statistical Analysis of a Cluster Randomized Controlled Trial

Xian Li[1,2*], MSc; Nadila Duolikun[1*], MPH; Fengzhuo Cheng[1], MPH; Laurent Billot[2,3], MRes; Weiping Jia[4,5], PhD; Puhong Zhang[1,2], PhD[‡]

[1]The George Institute for Global Health at Peking University Health Science Center, Beijing, China

[2]Faculty of Medicine, University of New South Wales, Sydney, Australia

[3]The George Institute for Global Health, Sydney, Australia

[4]Department of Endocrinology, Shanghai Jiaotong University Affiliated Sixth People's Hospital, Shanghai, China

[5]Chinese Diabetes Society, Beijing, China

[‡]ROADMAP study group

[*]these authors contributed equally

Corresponding Author:
Weiping Jia, PhD
Department of Endocrinology
Shanghai Jiaotong University Affiliated Sixth People's Hospital
600 Yishan Road
Shanghai
China
Phone: 86 21 64369181 ext 8922
Email: wpjia@sjtu.edu.cn

Abstract

Background: As the management of type 2 diabetes remains suboptimal in primary care, the Road to Hierarchical Diabetes Management at Primary Care (ROADMAP) study was designed and conducted in diverse primary care settings to test the effectiveness of a three-tiered diabetes management model of care in China.

Objective: This paper aims to predetermine the detailed analytical methods for the ROADMAP study before the database lock to reduce potential bias and facilitate transparent analyses.

Methods: The ROADMAP study adopts a community-based, cluster randomized controlled trial design that compares the effectiveness of a tiered diabetes management model on diabetes control with usual care among patients with diabetes over a 1-year study period. The primary outcome is the control rate of glycated hemoglobin (HbA_{1c}) <7% at 1 year. Secondary outcomes include the control rates of ABC (HbA_{1c}, blood pressure, and low-density lipoprotein cholesterol [LDL-C], individual and combined) and fasting blood glucose, and the change in each outcome. The primary analysis will be the log-binomial regression with generalized estimating equation (GEE), which accounts for the clustering within communities, for binary outcomes and linear regression with GEE for continuous outcomes. For both, the baseline value of the analyzed outcome will be the covariate. The other covariate further adjusted models and the repetitive models after multiple imputation (when more than 10% of observations in HbA_{1c} after 1 year are missing) will be used for sensitivity analysis. Five prespecified subgroup analyses have also been planned to explore the heterogeneity of the intervention effects by adding the subgroup variable and its interaction with the intervention to the primary model.

Results: This plan has been finalized, approved, and signed off by the principle investigator, co-principle investigator, and lead statisticians as of November 22, 2019, and made public on the institutional website without any knowledge of intervention allocation. Templates for the main figure and tables are presented.

Conclusions: This statistical analysis protocol was developed for the main results of the ROADMAP study by authors blinded to group allocation and with no access to study data, which will guarantee the transparency and reduce potential bias during statistical analysis.

Trial Registration: Chinese Clinical Trial Registry ChiCTR-IOC-17011325; https://tinyurl.com/ybpr9xrq

International Registered Report Identifier (IRRID): DERR1-10.2196/18333

(JMIR Res Protoc 2020;9(4):e18333) doi:10.2196/18333

KEYWORDS

statistical analysis plan; cluster randomized controlled trial; diabetes control; China; community-based; hemoglobin; primary care

Introduction

Study Background

Being home to the largest diabetic population, China has been encountering challenges in managing diabetes adequately in primary care [1-3]. Currently, patients with type 2 diabetes (T2D) are entitled to four or more yearly free blood glucose tests and treatment consultations at primary care clinics through a publicly funded essential public health service package [4]. Despite the provision of the universal access and increasing subsidies to community health care services [4,5], outcomes of the current management remain suboptimal. It has been reported that among service recipients only 40% have reached the adequate blood glucose control target (glycated hemoglobin [HbA_{1c}] <7.0%) [6], while less than a tenth have achieved optimal control of composite cardiometabolic "ABC" (HbA_{1c}, blood pressure [BP], and low-density lipoprotein cholesterol [LDL-C]) targets [7]. In July 2017, a cluster randomized controlled trial, Road to Hierarchical Diabetes Management at Primary Care (ROADMAP) study, was launched to determine the effectiveness of a strengthened version of the previously mentioned essential public health service on diabetes management through a mobile health–based, tiered service-delivery intervention in diverse primary care settings in China.

Study Overview

The ROADMAP study is designed as a community-based, cluster randomized controlled trial, aiming to compare the effectiveness of a hierarchical diabetes management intervention to usual care on blood glucose control. The usual care is the routine diabetes and hypertension management required by the national essential public health service. [4] The intervention delivery is performed by contracted service teams; each of them are composed of one primary care doctor at the community or village level as team leader, one township hospital doctor, and one district or county doctor. The intervention lasted for 1 year. This study is prospectively registered (ChiCTR-IOC-17011325) and ethically approved. A complete study description has been published elsewhere [8].

Participant Recruitment and Randomization

Participants are adult patients with established T2D who have registered for the essential public health service within the community at the time of recruitment. To be eligible, participants were 18-75 years old, resided in the community for the previous 6 months with no plan of relocating, and provided informed consent. Potential participants were excluded if they had severe physical or psychological injury or illness, were unable to attend the site visit or consciously answer questions, were women in the process of or planning for pregnancy or breastfeeding, or had participated in any other clinical trial within the previous 6 months.

The trial recruited a total of 19,601 participants from 864 communities or villages in 144 districts or counties in 25 provinces. Generally, for each participating province, 6 districts and 6 of its subordinate communities from each district were selected. The same principle applied to counties and villages in rural areas. Following the completion of baseline assessments, communities or villages (clusters) were centrally randomized.

Intervention

Besides a standard training workshop for the contracted service providers (community, township, and county level doctors) in the intervention arm, the key components of intervention were one BP measurement and two blood glucose monitoring tests (at least one fasting blood glucose [FBG]) monthly, instruction for lifestyle change and medication accordingly, timely referral if an indicator is present, and quarterly performance review for the contracted service team. A mobile health–based information system, Graded ROADMAP, was developed and employed to support the contracted doctor team delivering the intervention. Another smartphone app, Your Doctor, was available for participants in the intervention arm to facilitate health education and communication between the designated doctors and patients. The use of Your Doctor depended on participants' willingness and capability of using the smartphone app. At the end of this study, all the participants in the intervention arm were divided into two subgroups based on the actual use of Your Doctor: a basic intervention subgroup in which participants have logged in less than 4 times to the app throughout the 1-year follow-up and an intensive intervention subgroup in which the participants have at least 4 log-ins within 1 year.

Outcomes

The primary outcome is the HbA_{1c} control rate (percentage of patients achieving HbA_{1c} <7.0%; target A) at 1 year. The secondary outcomes include the percentage of patients achieving both systolic blood pressure <140 mmHg and diastolic blood pressure <80 mmHg (target B); the percentage of patients achieving LDL-C <2.6 mmol/L (target C); the optimal control rate of the combined ABC targets as previously defined; the percentage of patients achieving FBG <7.0 mmol/L; the changes in levels of HbA_{1c}, BP, LDL-C, and FBG; and the subtype-specific and overall hypoglycemia episodes [9]. Other outcomes are health-related quality of life measured by the EuroQol questionnaire EQ-5D-3L (3-level version of EQ-5D) [10,11], the mean change in the scores of the summary of diabetes self-care activities questionnaire [12], the development of any self-reported onset of new comorbidities and diabetic complications during follow-up, concomitant medications, and direct medical cost.

Sample Size

A sample size of 16,416 participants (19 patients per community) at 1 year provided an 89% power (2-sided α=.05) to detect a ≥5% absolute increase in the primary outcome for the intervention group. The sample size calculation assumes that 40% of participants will have well-controlled HbA_{1c} levels (<7%) at the end of the study in the control group [6], with an intraclass correlation coefficient of 0.15 based on our previous Observational Registry of Basal Insulin Treatment (ORBIT) study [13]. Furthermore, assuming that 50% of participants (ie, 576 clusters with a smaller average cluster size of 9-10 participants) in the intervention group will receive the intensive intervention (at the patients' discretion), it will need 93% to detect absolute increases of 5% HbA_{1c} control, when compared to the basic intervention group (576 clusters with an average cluster size of 9-10 patients). Accounting for a potential loss to follow-up of 14% of patients, the study aimed to recruit 19,008 patients with T2D from 864 communities or villages (576 in intervention and 288 in control, with a 2:1 ratio) in 24 provinces in mainland China, which equates to an average of 22 patients from each community or village.

Objectives

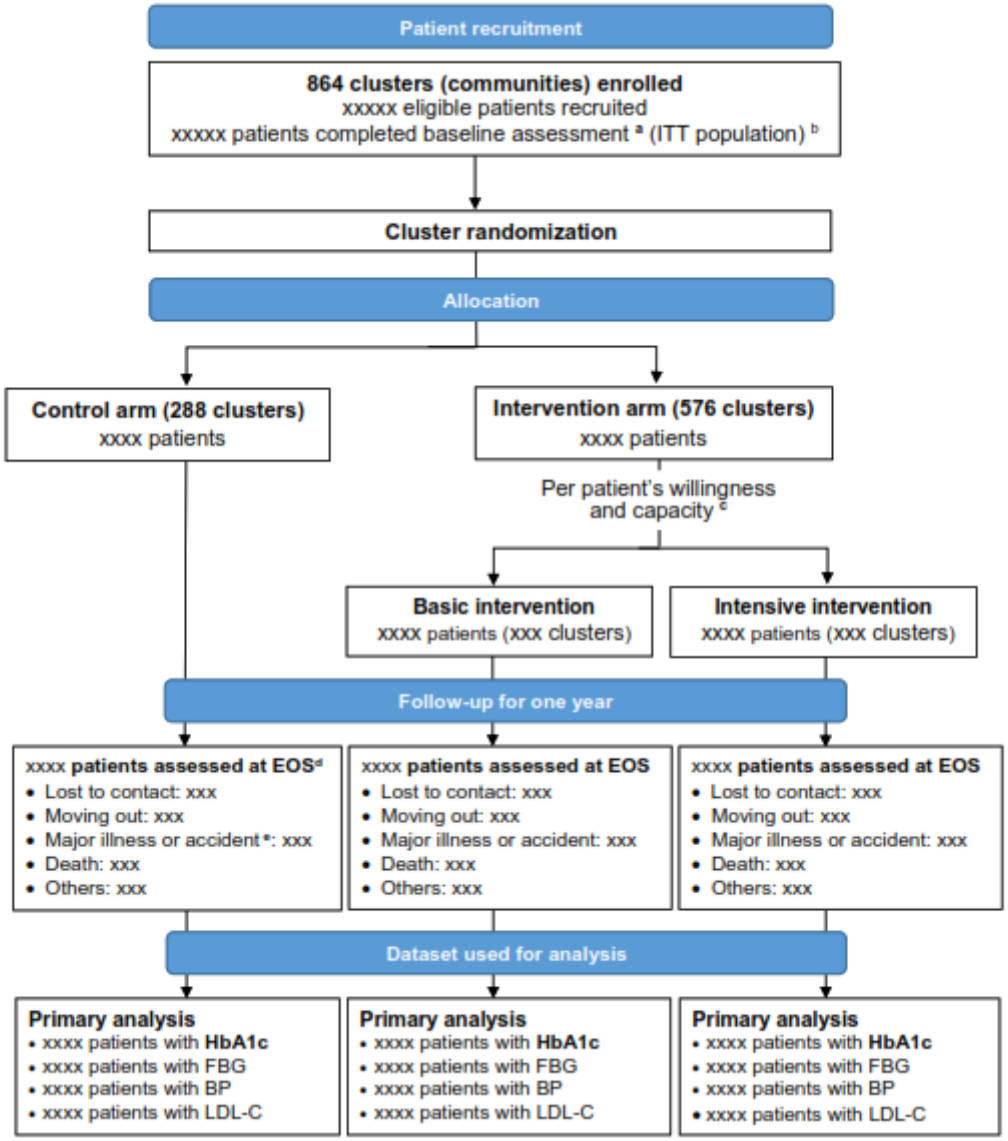
The purpose of this study is to outline the predetermined analytical methods in detail before completion of the database lock to reduce the potential bias and facilitate transparent analyses.

Methods

Patient Disposition

The flow of patients through the study will be displayed in a CONSORT (Consolidated Standards of Reporting Trials) diagram (figure shell is shown in Figure 1).

Figure 1. CONSORT diagram for the randomized control trial. BP: blood pressure; EOS: end of study; FBG: fasting blood glucose; HbA_{1c}: glycated hemoglobin; ITT: intention-to-treat; LDL-C: low-density lipoprotein cholesterol.



Population for Analysis

The main analysis will be performed at the patient level following an intention-to-treat (ITT) principle (ie, patients and clusters will be included for analysis as their assigned group, regardless of treatment adherence). All randomized patients that have given consent and are not missing key variables (including gender, age, and HbA_{1c} level) at the baseline assessment will form the ITT population.

General Analysis Principles

Analyses will adjust for clustering at the village level. The intracluster correlation coefficients will be calculated and tabulated. The primary comparison will be made between the intervention population and the usual care population. As a secondary comparison, the effects of the intensive intervention (no less than 4 logins to the Your Doctor app in addition to the basic intervention) will be explored by estimating the effects of intensive intervention vs basic intervention. No formal interim analysis will be performed. The primary analysis will use all available data with no imputation. Imputation will be performed as a sensitivity analysis when missing observations in HbA_{1c} levels after 1 year exceed 10%.

Hypothesis tests will be 2-tailed with a 5% significance level maintained throughout the analyses. No adjustment will be applied for multiplicity, given that most of the effectiveness outcomes are correlated or consist in different versions of common variables. Outcomes will be presented in order of their priority (primary vs secondary), and only a limited number of subgroup analyses are prespecified. Subgroup analyses will be carried out irrespective of whether a significant treatment effect on the primary outcome is observed. Analyses will be performed in SAS, version 9.4 or later (SAS Institute).

Patient Characteristics and Baseline Comparison

Discrete baseline variables will be summarized by frequencies and percentages. Percentages will be calculated using the number of patients for whom data is available. Continuous variables will be summarized by using mean and SD, and median and interquartile range (Q1-Q3). No adjustment for clustering will be applied when comparing baseline characteristics. Standardized differences between groups in baseline characteristics will be reported. A less than 0.1 standard difference will be used to indicate a negligible difference in the mean or proportion of a baseline variable between treatment groups [14].

Variables for baseline measures will be tabulated, including demographic (age, sex, ethnicity), socioeconomic status (education level, household income, health insurance), anthropometric measurements (weight, BMI, waist circumference, BP), smoking status, diabetic complications, laboratory data (HbA_{1c}, FBG, LDL-C, serum creatinine), and score of the summary of diabetes self-care activities.

Cluster characteristics will be summarized in stratified and overall randomization groups. The cluster characteristics at baseline are age of the community or village doctors, smartphone operation system (iOS or Android) that the community or village doctors were using, and number of patients with T2D registered.

Compliance to Basic Intervention Service

Data on the receipt of intervention provided for standardized diabetes management services have been routinely collected by the app Graded ROADMAP, including the frequencies and the results of FBG or postprandial blood glucose tests, BP measurements, and patient referrals from primary care clinics to the upstream hospitals at the county or district level. According to the protocol, each patient is due to receive at least two blood glucose tests (FBG, postprandial blood glucose, or both) and one BP measurement per month. Although necessary referral is encouraged when indicators are present, there is no requirement for the referral rate.

Among the intervention group, the mean times of health services received per patient per month; the percentage of patients achieving the protocol-required measurement frequency; and the mean values of FBG, postprandial blood glucose, and BP will be illustrated in separate figures with a double y-axis.

Exposure to Intensive Intervention

Within the intervention group, the frequency of patient's log-ins to the Your Doctor app will be described in numbers and percentages of frequency from 0 to >12 over the 1-year study period, and as a categorical variable in frequency ≥4 or <4. A patient with the frequency ≥4 times per year (referring to the essential public health service recommendation) will be defined as a complier to intensive intervention and, thus, form the intensive intervention subset for further comparison between basic and intensive intervention.

Analysis of the Primary Outcome

Primary Analyses

The primary endpoint, control rate of HbA_{1c} at 1 year, will be first compared between all intervention groups and all control groups. The primary analysis of the intervention effect will be conducted using a log-binomial regression with generalized estimating equations (GEE) to account for clustering within communities and with adjustment for baseline HbA_{1c} as a continuous covariate (model 1). The raw number and percentage of patients with adequate control of HbA_{1c} at 1 year will be reported. The effect of the intervention will be presented as the relative risk (RR) of the proportion of HbA_{1c} levels <7% along with its 95% CI and corresponding *P* value. In cases of a convergence issue, the logistic regression with GEE and the Poisson regression with GEE will be both used as the alternative methods, with the Poisson regression as the sensitivity analysis. The odds ratio along with the indirectly derived RR will be reported for the logistic regression.

Covariates-Adjusted Analyses

The primary model (model 1) previously described will be rerun after further adjusting for the following covariates (model 2): economic development level; locality; and other baseline covariates that have shown a significant difference ($P<.01$) between the intervention and control groups in a univariate

comparison, including all the baseline variables previously listed (see "Patient Characteristics and Baseline Comparison").

Imputed Data Analyses

Multiple imputation using fully conditional specification [15] will be performed as a sensitivity analysis when more than 10% of observations in HbA_{1c} is missing at 1 year. The imputation model will include the levels of HbA_{1c}, FBG, BP, and LDL-C at 1 year and at baseline; a cluster indicator; a group indicator; and other baseline variables including age, sex, education, duration of diabetes, comorbidities, economic development level, and locality (urban or rural). Ten sets of imputed data will be created and analyzed using the model 1 (described in "Primary Analyses"). HbA_{1c}, FBG, BP, and LDL-C will first be imputed as continuous variables using linear regression and subsequently converted into binary variables. Estimates of the intervention effect after imputation (β in model 1) and its standard errors will be combined to obtain a pooled common RR and 95% CI.

Subgroup Analyses

Five prespecified subgroup analyses will be carried out between the overall intervention group and the control group. The subgroups are economic development level (developed vs less-developed), locality (urban vs rural), age group (<60 years of age vs ≥60 years of age), duration of diabetes (≥6 years vs <6 years, around the median), and diabetic complication (yes vs no; yes is defined as presence of any diagnosed diabetic nephropathy, diabetic retinopathy, peripheral neuropathy, carotid artery disease, lower extremity artery disease, diabetic foot damage, peripheral vascular disease, coronary stenosis, myocardial infarction, postcoronary artery surgery, cerebral infarction, or cerebral hemorrhage).

The analysis for each subgroup analysis will be performed by adding the subgroup variable along with its interaction with the intervention as fixed effects to the primary model (model 1). Within each subgroup, the raw counts and percentages within each treatment arm will be presented, as well as the RRs and their 95% CIs for the intervention effect from the primary model. The results will be displayed on a forest plot including the *P* value for heterogeneity corresponding to the interaction term between the intervention and the subgroup variable.

Comparison Between Basic and Intensive Intervention

To explore the possible additional effect from intensive intervention, we will compare the intensive intervention to the basic intervention using similar models, regardless of whether the difference between the overall intervention group and control group is statistically significant. Considering the potential imbalance in the baseline characteristics between patients following the intensive intervention and those following the basic intervention, two propensity score methods will be applied. Both the propensity score adjusted regression based on a log-binomial model with GEE and the inverse probability of treatment weighting method will be used to evaluate the intervention effect [16,17] as the sensitivity to seek the consistency of the conclusion. The propensity score model will consist in a simple logistic regression with baseline covariates including age, sex, the baseline value of the analyzed outcome, education, duration of diabetes, comorbidities, economic development level, and locality (urban or rural). Baseline characteristics of participants will be described in a table separately before and after propensity score adjustments.

Analysis of Secondary Outcomes

Analysis for Binary Categorical Outcomes

A similar analytic strategy as for the primary outcome will be followed for other binary outcomes. These include the proportion of FBG<7.0 mmol/L, BP<140/80 mmHg, and that of LDL-C<2.6 mmol/L, and the optimal control rate of combined ABC targets. A log-binomial regression with GEE and including baseline continuous values of the analyzed outcome variable as covariate (model 1) will be used. Other further adjusted analysis (model 2) and subgroup analysis will also be conducted, and the imputed analysis as well, if applicable.

Analysis for Continuous Outcomes

HbA_{1c}, FBG, BP, and LDL-C will also be analyzed as continuous variables. A similar analytic strategy as the one used for binary outcomes will be followed but using a linear regression (ie, assuming a normal distribution and an identity link, instead of a log-binomial regression). The raw mean (SD) of the changes will be reported. The effect of the intervention will be presented as the adjusted mean difference and associated 95% CIs. Further covariates-adjusted analyses (see model 2) and subgroup analysis will also be conducted, and the imputed analysis as well, if applicable.

Analysis of Hypoglycemia

Episodes of hypoglycemia (each subtype and overall) will be analyzed with the same approach as before; this time using Poisson regression adjusted for the baseline count of hypoglycemia. The effect of the intervention will be estimated as the incidence rate of hypoglycemia episodes and its 95% CI. The number of patients experiencing at least one hypoglycemia episode and the total number of episodes will be tabulated by group. Adjusted analysis and subgroup analysis will also be conducted. No imputation will be performed on hypoglycemia.

Analysis of Other Outcomes

No subgroup and imputed analysis will be performed on the following endpoints.

EQ-5D

The EQ-5D index value and EuroQol-visual analogue scale (EQ-VAS) score at baseline and at end-of-study will be described using mean and SD by treatment groups. For changes in EQ-5D index values and VAS scores, a linear regression with GEE accounting for clustering will be used to test the difference between groups. The baseline values of the outcome variable will be included as covariates. The raw mean (SD) of the score changes will be reported by treatment groups; the effect of the intervention will be presented as the mean differences of the changes and associated 95% CIs.

The Summary of Diabetes Self-Care Activities

The scores of the summary of diabetes self-care activities questionnaire will be described at baseline and 1 year. For their

changes from baseline, a linear regression with GEE and with adjustment of the baseline values of analyzed outcome variables will be used to compare the difference between treatment groups. Other covariates with further adjusted models will also be conducted. The raw mean (SD) of the score changes will be reported by treatment groups; the effect of the intervention will be presented as the mean differences and associated 95% CIs from the two previously mentioned models.

Other Variables

The following outcomes will be analyzed using descriptive statistics without any adjustment for clustering.

1. Concomitant medications. Insulin injection and oral antidiabetic drug intake at baseline and at end-of-study assessments. The combination of different oral antidiabetic drug regimens or the combination of insulin injection and oral antidiabetic drugs, or their single use will also be summarized.
2. New-onset comorbidities and diabetic complications will be described as the numbers and percentages of patients with each new onset complication by treatment groups.
3. Direct medical cost is the self-reported direct cost on medication and medical expense for health care services (inpatient or outpatient cost, medication cost), including total cost and out-of-pocket cost.

Results

This study was funded in January 2017. Ethics approval was obtained from the Institutional Review Board at Shanghai Sixth People's Hospital, where the lead PI is affiliated with, before the study commenced. Written approval from each participating site was granted by the local hospital research ethics committee, and other relevant regional regulatory bodies. Signed informed consent was obtained from all trial participating doctors and patients prior to participant recruitment. Recruitment commenced in June 2017 and closed after completed baseline assessment in December 2018 for all 864 trial participating communities in 144 districts or counties in 25 provincial sites (1 more province than the scheduled 24 because of a shortage in eligible district or county hospitals). As of October 2019, the last 1-year end-of-study assessment ended. The internal statistical plan was reviewed, approved, and signed off in November 2019 and made public on the institutional internal website prior to the database lock in January 2020.

Templates of main tables (ie, baseline characteristics as in Table 1, estimated intervention effects for binary outcomes as in Table 2, estimated intervention effects for continuous outcomes as in Table 3, and hypoglycemia incidence as in Table 4) were produced prior to the previously mentioned analytical methods.

Table 1. Baseline characteristics of participants by treatment arms in the road to hierarchical diabetes management at primary care study.

Characteristics	Control (xxxx)	Intervention (xxxx)	Standardized differences
Region by economic development, n (%)			x.xxx
Developed	xxx (xx.x)	xxx (xx.x)	
Less developed	xxx (xx.x)	xxx (xx.x)	
Locality, n (%)			x.xxx
Urban	xxx (xx.x)	xxx (xx.x)	
Rural	xxx (xx.x)	xxx (xx.x)	
Demographics			
Age (years), mean (SD)	xx (xx)	xx (xx)	x.xxx
Gender (male), n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
BMI (kg/m^2), mean (SD)	xxx (xxx)	xxx (xxx)	x.xxx
Highest level of education, n (%)			x.xxx
Primary school or lower	xxx (xx.x)	xxx (xx.x)	
Junior high school	xxx (xx.x)	xxx (xx.x)	
Senior high school	xxx (xx.x)	xxx (xx.x)	
Junior college and above	xxx (xx.x)	xxx (xx.x)	
Annual income per capita (CNY), mean (SD)	xxx (xxx)	xxx (xxx)	x.xxx
Health insurance coverage (Yes), n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Insurance reimbursement rates, n (%)			x.xxx
70%-100%	xxx (xx.x)	xxx (xx.x)	
50%-70%	xxx (xx.x)	xxx (xx.x)	
<50%	xxx (xx.x)	xxx (xx.x)	
Self-reported medical history and complications			
Duration of diabetes (years), median (Q1, Q3)	xxx (xxx, xxx)	xxx (xxx, xxx)	x.xxx
Current smoker, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Hypertension, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Dyslipidemia, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Diabetic nephropathy, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Diabetic retinopathy, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Peripheral neuropathy, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Lower extremity, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Macro-vascular, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Lab characteristics			
HbA_{1c}[a] (%), mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xxx
HbA_{1c}<7%, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
FBG[b] (mmol/L), mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xxx
FBG<7.0 mmol/L, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
SBP[c] (mmHg), mean (SD)	xxx (x.xx)	xxx (x.xx)	x.xxx
DBP[d] (mmHg), mean (SD)	xxx (x.xx)	xxx (x.xx)	x.xxx
LDL-C[e] (mmol/L), mean (SD)	x.xx (xxx)	x.xx (xxx)	x.xxx
LDL-C<2.6 mmol/L, n (%)	xxx (xx.x)	xxx (xx.x)	x.xxx
Serum creatinine (umol/L), median (Q1, Q3)	xxx (xxx, xxx)	xxx (xxx, xxx)	x.xxx

[a]HbA_{1c}: glycated hemoglobin.

[b]FBG: fasting blood glucose.

[c]SBP: systolic blood pressure.

[d]DBP: diastolic blood pressure.

[e]LDL-C: low-density lipoprotein cholesterol.

Table 2. Estimated effects of intervention compared to control on primary and secondary binary outcomes at end of study.

Outcomes	Control, n (%)	Intervention, n (%)	Primary model[a]	
			RR[b] (95% CI)	*P* value
Primary outcome				
HbA_{1c}[c] < 7.0%	xxx (xx.x)	xxx (xx.x)	x.xx (x.xx-x.xx)	.xx
Secondary outcomes				
FBG[d] < 7.0 mmol/L	xxx (xx.x)	xxx (xx.x)	x.xx (x.xx-x.xx)	.xx
BP[e] < 140/80 mmHg[f]	xxx (xx.x)	xxx (xx.x)	x.xx (x.xx-x.xx)	.xx
LDL-C[g] < 2.6 mmol/L	xxx (xx.x)	xxx (xx.x)	x.xx (x.xx-x.xx)	.xx
Composite diabetes control[h,i]	xxx (xx.x)	xxx (xx.x)	x.xx (x.xx-x.xx)	.xx

[a]Primary model: log-binomial regression with generalized estimating equation (GEE) with adjustment of the baseline value of the analyzed outcome and clustering. The logistic regression with GEE will be employed as the alternative method in case of non-convergence, with indirectly derived relative risk reported.

[b]RR: relative risk.

[c]HbA_{1c}: glycated hemoglobin.

[d]FBG: fasting blood glucose.

[e]BP: blood pressure.

[f]Only systolic blood pressure at baseline and clustering were adjusted in the primary model for BP control.

[g]LDL-C: low-density lipoprotein cholesterol.

[h]Composite diabetes control: defined as HbA_{1c} level <7.0%, BP <140/80 mmHg and LDL-C <2.6 mmol/L.

[i]No baseline variable was adjusted in the primary model for the composite diabetes control.

Table 3. Estimated effects of intervention compared to control on the change from baseline of continuous outcomes.

Secondary continuous outcome	Control, mean (SD)	Intervention, mean (SD)	Primary model[a]	
			Mean differences (95% CI)	*P* value
The change from baseline of:				
HbA_{1c}[b] level, %	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx-x.xx)	.xx
FBG[c] level, mmol/L	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx-x.xx)	.xx
Systolic blood pressure, mmHg	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx-x.xx)	.xx
Diastolic blood pressure, mmHg	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx-x.xx)	.xx
LDL-C[d] level, mmol/L	x.xx (x.xx)	x.xx (x.xx)	x.xx (x.xx-x.xx)	.xx

[a]Primary model: linear regression with generalized estimating equation and with adjustment of baseline value of the analyzed outcome and clustering.

[b]HbA_{1c}: glycated hemoglobin.

[c]FBG: fasting blood glucose.

[d]LDL-C: low-density lipoprotein cholesterol.

Table 4. Incidence and events rate of hypoglycemia by treatment arms.

Hypoglycemia[a]	Control (n=xxxx)			Intervention (n=xxxx)			*P* value[b]
	Patients	Events	Events/100 patients	Patients	Events	Events/100 patients	
Symptomatic hypoglycemia	xxxx	xxxx	xx	xxxx	xxxx	xx	.xx
Asymptomatic hypoglycemia	xxxx	xxxx	xx	xxxx	xxxx	xx	.xx
Probable symptomatic hypoglycemia	xxxx	xxxx	xx	xxxx	xxxx	xx	.xx
Relative hypoglycemia	xxxx	xxxx	xx	xxxx	xxxx	xx	.xx
Overall hypoglycemia	xxxx	xxxx	xx	xxxx	xxxx	xx	.xx

[a]Hypoglycemia subtypes followed the American Diabetes Association and The Endocrine Society suggested classifications [9].

[b]*P* values will be reported from Poisson regression with generalized estimating equation and with adjustment of baseline count of each hypoglycemia category.

Discussion

Summaries

This article presents the detailed statistical analysis plan for the ROADMAP study, which is a clustered randomized controlled trial conducted in diverse areas of China with the purpose of testing the effectiveness of a mobile health platform named *Graded Roadmap* on diabetes control.

The clustered randomized controlled trial design is useful for assessing community-based interventions like the ROADMAP study yet requires careful attention to conduct valid analyses. As such, clustering of outcomes was accounted for when designing the study and, as previously mentioned, will be accounted for in the analysis of the study outcomes. In the study design, the intraclass correlation coefficient was estimated to be 0.15 based on our previous ORBIT study. ORBIT recruited patients who were initiating basal insulin treatment at secondary and tertiary hospitals. The intraclass correlation coefficient may differ from that observed in the studied populations in ROADMAP.

To find out the effectiveness and feasibility of the platform in different regions, four strata were covered by ROADMAP. They are economically developed urban areas, economically developed rural areas, economically less-developed urban areas, and economically less-developed rural areas. Although the economic development level (developed vs less-developed) and locality (urban vs rural) have been included in the predetermined subgroup analysis, more detailed descriptions might be needed for the factorial subgroups.

Conclusions

This statistical analysis plan was developed for the main results of the ROADMAP study by authors blinded to group allocation and with no access to study data, which will guarantee the transparency and reduce potential bias during statistical analysis.

Acknowledgments

This study received its principle funding from DongBao Pharmaceutical Co, Ltd through Bethune Charitable Foundation. DongBao Pharmaceutical Co, Ltd also provided in-kind support including the decision-making system, blood glucose monitoring equipment (smart blood glucose meters), and consumables to facilitate the implementation of the intervention. Wuxi BioHermes Bio&Medical Technology Co, Ltd provided point-of-care HbA_{1c} analyzers (A1c EZ 2.0) for universal testing of HbA_{1c}. None of the funders participated in the study design, implementation, data analysis and interpretation of the results.

Authors' Contributions

XL and ND led the manuscript drafting and XL led the development of the methods for statistical analysis. WJ and PZ equally conceived the project and designed the main study. WJ, PZ, ND, and XL contributed to the statistical design of the study. PZ, ND, FC, and LB participated in the design of the statistical analysis methods. All authors reviewed and contributed to the refinement of the manuscript and approved the submitted final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 10622 KB - resprot_v9i4e18333_app1.pdf]

References

1. International Diabetes Federation. 2019. IDF diabetes atlas, 9th edition URL: http://www.diabetesatlas.org/ [accessed 2020-01-20]
2. Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 Sep 28;394(10204):1145-1158. [doi: 10.1016/S0140-6736(19)30427-1] [Medline: 31248666]
3. Mao W, Yip CW, Chen W. Complications of diabetes in China: health system and economic implications. BMC Public Health 2019 Mar 06;19(1):269. [doi: 10.1186/s12889-019-6569-8] [Medline: 30841928]
4. National Health and Family Planning Commission, P.R.China. 2017. Notice of National Health and Family Planning Commission on Printing and Issuing the Essential Public Health Services Specificaiton (3rd edition) URL: http://www.nhc.gov.cn/jws/s3578/201703/d20c37e23e1f4c7db7b8e25f34473e1b.shtml [accessed 2020-01-20]
5. Ministry of Finance, P.R. China. 2019. Report on the Execution of the Central and Local Budgets for 2018 and on the Draft Central and Local Budgets for 2019 URL: http://www.xinhuanet.com/english/2019-03/17/c_137901687.htm [accessed 2020-01-20]
6. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, 2010 China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. JAMA 2013 Sep 04;310(9):948-959. [doi: 10.1001/jama.2013.168118] [Medline: 24002281]
7. Ji L, Hu D, Pan C, Weng J, Huo Y, Ma C, CCMR Advisory Board, CCMR-3B STUDY Investigators. Primacy of the 3B approach to control risk factors for cardiovascular disease in type 2 diabetes patients. Am J Med 2013 Oct;126(10):925.e11-925.e22. [doi: 10.1016/j.amjmed.2013.02.035] [Medline: 23810406]
8. Jia W, Zhang P, Duolikun N, Zhu D, Li H, Bao Y, ROADMAP study group. Study protocol for the road to hierarchical diabetes management at primary care (ROADMAP) study in China: a cluster randomised controlled trial. BMJ Open 2020 Jan 06;10(1):e032734. [doi: 10.1136/bmjopen-2019-032734] [Medline: 31911516]
9. Workgroup on Hypoglycemia, American Diabetes Association. Defining and reporting hypoglycemia in diabetes: a report from the American Diabetes Association Workgroup on Hypoglycemia. Diabetes Care 2005 May;28(5):1245-1249. [Medline: 15855602]
10. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
11. Liu GG, Wu H, Li M, Gao C, Luo N. Chinese time trade-off values for EQ-5D health states. Value Health 2014 Jul;17(5):597-604. [doi: 10.1016/j.jval.2014.05.007] [Medline: 25128053]
12. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [Medline: 10895844]
13. Ji L, Zhang P, Weng J, Lu J, Guo X, Jia W, et al. Observational Registry of Basal Insulin Treatment (ORBIT) in patients with type 2 diabetes uncontrolled by oral hypoglycemic agents in China--study design and baseline characteristics. Diabetes Technol Ther 2015 Oct;17(10):735-744. [doi: 10.1089/dia.2015.0054] [Medline: 26171728]
14. Normand ST, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. J Clin Epidemiol 2001 May;54(4):387-398. [doi: 10.1016/s0895-4356(00)00321-8] [Medline: 11297888]
15. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. Stat Methods Med Res 2007 Jun;16(3):219-242. [doi: 10.1177/0962280206074463] [Medline: 17621469]
16. Vansteelandt S, Daniel RM. On regression adjustment for the propensity score. Stat Med 2014 Oct 15;33(23):4053-4072. [doi: 10.1002/sim.6207] [Medline: 24825821]
17. Benedetto U, Head SJ, Angelini GD, Blackstone EH. Statistical primer: propensity score matching and its alternatives. Eur J Cardiothorac Surg 2018 Jun 01;53(6):1112-1117. [doi: 10.1093/ejcts/ezy167] [Medline: 29684154]

Abbreviations

ABC: HbA_{1c}, BP, and LDL-C
BP: blood pressure
CONSORT: Consolidated Standards of Reporting Trials
DBP: diastolic blood pressure
EQ-5D-3L: 3-level version of EQ-5D
EQ-VAS: EuroQol-visual analogue scale
FBG: fasting blood glucose
GEE: generalized estimating equation
HbA_{1c}: glycated hemoglobin
ITT: intention-to-treat
LDL-C: low-density lipoprotein cholesterol
ORBIT: Observational Registry of Basal Insulin Treatment
ROADMAP: Road to Hierarchical Diabetes Management at Primary Care

RR: relative risk
T2D: type 2 diabetes

Edited by G Eysenbach; submitted 20.02.20; peer-reviewed by P Gao, W Zhang; comments to author 07.03.20; revised version received 10.03.20; accepted 22.03.20; published 28.04.20.

Please cite as:
Li X, Duolikun N, Cheng F, Billot L, Jia W, Zhang P
Road to Hierarchical Diabetes Management at Primary Care (ROADMAP) Study in China: Protocol for the Statistical Analysis of a Cluster Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e18333
URL: http://www.researchprotocols.org/2020/4/e18333/
doi:10.2196/18333
PMID:32343257

Protocol

Effect of Various Invitation Schemes on the Use of Fecal Immunochemical Tests for Colorectal Cancer Screening: Protocol for a Randomized Controlled Trial

Laura Fiona Gruner[1,2], MSc; Michael Hoffmeister[1], PhD; Leopold Ludwig[3], MD; Hermann Brenner[1,4,5], MD, MPH

[1]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany

[2]Medical Faculty Heidelberg, University of Heidelberg, Heidelberg, Germany

[3]Gastroentereologische Schwerpunktpraxis, Dornstadt, Germany

[4]Division of Preventive Oncology, German Cancer Research Center and National Center for Tumor Diseases, Heidelberg, Germany

[5]German Cancer Consortium, German Cancer Research Center, Heidelberg, Germany

Corresponding Author:
Hermann Brenner, MD, MPH
Division of Clinical Epidemiology and Aging Research
German Cancer Research Center
Im Neuenheimer Feld 581
Heidelberg, 69120
Germany
Phone: 49 6221 42 1300
Email: h.brenner@dkfz.de

Abstract

Background: Fecal occult blood testing has been offered for many years in the German health care system, but participation rates have been notoriously low.

Objective: The aim of this study is to evaluate the effect of various personal invitation schemes on the use of fecal immunochemical tests (FITs) in persons aged 50-54 years.

Methods: This study consists of a three-armed randomized controlled trial: (1) arm A: an invitation letter from a health insurance plan including a FIT test kit, (2) arm B: an invitation letter from a health insurance plan including an offer to receive a free FIT test kit by mail upon easy-to-handle request (ie, by internet, fax, or reply mail), and (3) arm C: an information letter on an existing colonoscopy offer (ie, control). Within arms A and B, a random selection of 50% of the study population will receive reminder letters, the effects of which are to be evaluated in a substudy.

Results: A total of 17,532 persons aged 50-54 years in a statutory health insurance plan in the southwest of Germany—AOK Baden-Wuerttemberg—were sent an initial invitation, and 5825 reminder letters were sent out. The primary end point is FIT usage within 1 year from receipt of invitation or information letter. The main secondary end points include gender-specific FIT usage within 1 year, rates of positive test results, rates of colonoscopies following a positive test result, and detection rates of advanced neoplasms. The study was launched in September 2017. Data collection and workup were completed in fall 2019.

Conclusions: This randomized controlled trial will provide important empirical evidence for enhancing colorectal cancer screening offers in the German health care system.

Trial Registration: German Clinical Trials Register (DRKS) DRKS00011858; https://bit.ly/2UBTIdt

International Registered Report Identifier (IRRID): DERR1-10.2196/16413

(JMIR Res Protoc 2020;9(4):e16413) doi:10.2196/16413

KEYWORDS

colorectal cancer; early detection; screening; fecal immunochemical test (FIT); invitation; Germany

Introduction

Background

Annual or biennial screening by fecal occult blood tests (FOBTs) has been shown by randomized clinical trials to reduce colorectal cancer (CRC) mortality by up to 30% [1]. These effects were achieved even with guaiac-based FOBTs (gFOBTs), which had limited sensitivity in detecting CRC and its precursors. Even stronger effects are to be expected by screening with immunochemical FOBTs (iFOBTs), often called fecal immunochemical tests (FITs), which have substantially higher sensitivity than gFOBTs [2]. FITs are now commonly recommended for CRC screening by national and international guidelines [2-4], and are increasingly offered for CRC screening in many countries [5,6]. However, participation rates in screening have remained low in countries where screening is offered in an opportunistic manner without targeted invitation of the eligible population [5]. This particularly applies to Germany, where gFOBT screening had been offered from 1977 to March 2017; FIT-based screening has been offered from age 50 years on since April 2017. For conducting FITs that are covered by the health insurance system, people have to pick up and return the tests at medical practices. Although personal information letters on CRC screening offers have been sent to the eligible population since July 2019 [3], they do not include FITs or specific low-threshold access to FITs, which are deemed to be crucial to achieve high participation rates [7,8]. The aim of this trial is to assess the effect of various invitation schemes on use of FITs for CRC screening in routine practice in the German health care system.

Objectives

The primary outcome that will be investigated is as follows: determine the proportion of people completing a FIT within 1 year after receiving a personal invitation letter within each trial arm.

Secondary outcomes are to determine the following:

1. The rate of positive test results.
2. The rate of performing a colonoscopy after a positive test.
3. The rate of performing a colonoscopy after a negative test.
4. The rate of discovered advanced colorectal neoplasia (ie, advanced adenomas and cancer).
5. The positive predictive value of the test.
6. The rate of performing a screening colonoscopy, in general, within 1 year.

Gender-specific analyses will be conducted.

Methods

Setting and Design

This protocol follows the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2013 Statement for clinical trial protocols [9] and the Fecal Immunochemical TesTs for Hemoglobin Evaluation Reporting (FITTER) guidelines [10].

The study is being conducted by the German Cancer Research Center (DKFZ) in cooperation with a statutory health insurance plan in the southwest of Germany: AOK Baden-Wuerttemberg (AOK BW).

The study is designed as a three-armed randomized controlled trial, in which 17,532 people aged 50-54 years are randomly selected to receive a letter from their health insurance plan (ie, AOK BW) with the following: (1) arm A: an invitation letter including a FIT test kit, (2) arm B: an invitation letter including an offer to receive a free FIT test kit by mail upon easy-to-handle request (ie, via Internet, fax, or reply mail), or (3) arm C: an information letter on CRC screening in the form of a colonoscopy; the control group represents routine practice with no study-related adaption of the letter. There are about 5844 insurants per arm. In addition, a reminder letter is being sent by random selection to 50% of the population in arms A and B. Prior to the recruitment of participants, the study, which was launched in 2017, was approved by the ethics committee of the Medical Faculty of the University of Heidelberg, Germany. It was registered in the German Clinical Trials Register (DRKS) on March 20, 2017 (DRKS-ID: DRKS00011858) with the following title: *Increase of the usage and effectiveness of colorectal cancer screening by means of targeted invitations with and without providing fecal immunochemical tests*. Written informed consent is being obtained from all participants.

Regardless of the FITs offered to participants in arms A and B within the intervention groups, the entire targeted population has access to CRC screening offers in routine practice according to law. However, only the opportunistic screening program is available at the time of study recruitment: CRC screening comprises annual testing for blood in the stool using the FIT or by performing a screening colonoscopy from age 50 years for men and women within a specific health insurance plan of AOK BW called *AOK-FacharztProgramm* (AOK Specialist Program); screening colonoscopy is otherwise offered from age 55 years in Germany during the recruitment period.

The study consists of two parts: (1) the mailing of different personal invitation letters for CRC screening, with low-threshold access to a FIT in the intervention arms, and (2) the follow-up of the use and outcome of a colonoscopy after a positive FIT, as well as the assessment of conducted FITs and colonoscopies in routine practice within 1 year after the initial invitation. Routine FIT and colonoscopy usage is derived from AOK BW claims data, which is being aggregated and pseudonymized where consent was obtained. All collected information is being stored and monitored in a study database by trained staff. The study design and the study assessments are shown in Figure 1 and Table 1, respectively.

Figure 1. Study design. Total expected fecal immunochemical test (FIT) use per study arm: A=35%, B=17.5%, and C=10%. Participants in arms A and B may receive a FIT as part of the intervention or can additionally request it in routine practice. Participants in arm C (ie, control) can only use the FIT in routine practice. Data on the use of screening methods until 1 year after the initial invitation will be collected. AOK BW: AOK Baden-Wuerttemberg; A1 and A2: subgroups of arm A; and B1 and B2: subgroups of arm B.

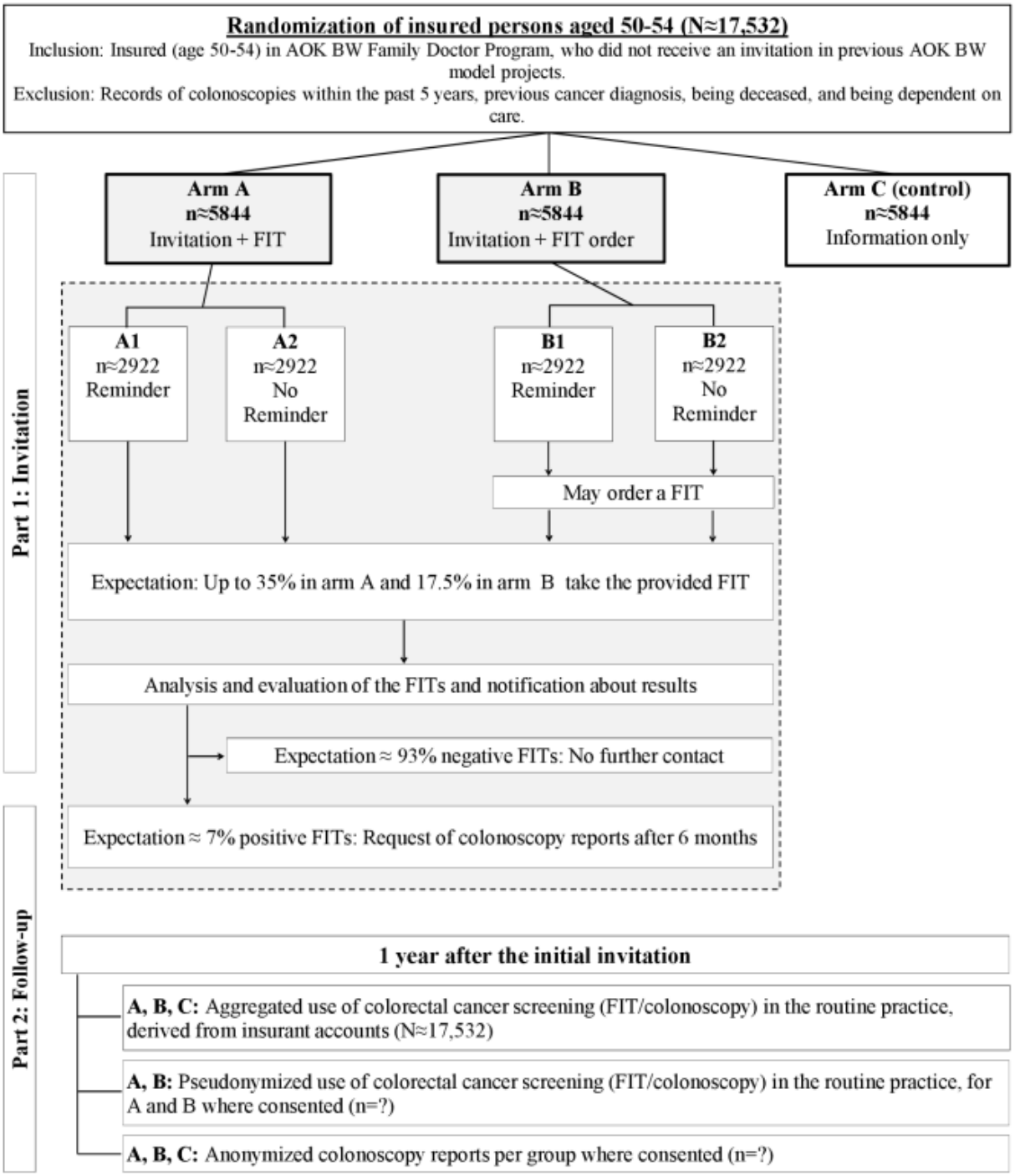


Table 1. Overview of study assessments at relevant time points.

Assessments	T0 (0 wk)	T1								T2 (6 months after positive FIT[a])	T3 (1 year)
		2 wk	4 wk	6 wk	8 wk	10 wk	12 wk	14 wk	>14[b] wk		
Randomization and initial invitations											
Randomization[c]	✓	_[d]	–	–	–	–	–	–	–	–	–
Subrandomization[e]	✓	–	–	–	–	–	–	–	–	–	–
Tranche 1	✓	–	–	–	–	–	–	–	–	–	–
Reminder 1	–	–	✓	–	–	–	–	–	–	–	–
Tranche 2	–	✓	–	–	–	–	–	–	–	–	–
Reminder 2	–	–	–	✓	–	–	–	–	–	–	–
Tranche 3	–	–	✓	–	–	–	–	–	–	–	–
Reminder 3	–	–	–	–	✓	–	–	–	–	–	–
Tranche 4	–	–	–	✓	–	–	–	–	–	–	–
Reminder 4	–	–	–	–	–	✓	–	–	–	–	–
Tranche 5	–	–	–	–	✓	–	–	–	–	–	–
Reminder 5	–	–	–	–	–	–	✓	–	–	–	–
Tranche 6	–	–	–	–	–	✓	–	–	–	–	–
Reminder 6	–	–	–	–	–	–	–	✓	–	–	–
Laboratory analysis of the provided FITs and notification about results											
Laboratory analysis	–	✓	✓	✓	✓	✓	✓	✓	✓	–	–
Notification	–	✓	✓	✓	✓	✓	✓	✓	✓	–	–
Follow-up of colonoscopy use after a positive FIT: request reports	–	–	–	–	–	–	–	–	–	✓	–
Use of colorectal cancer screening in routine practice within 1 year											
Use of FITs[f]	–	–	–	–	–	–	–	–	–	–	✓
Use of colonoscopy[f]	–	–	–	–	–	–	–	–	–	–	✓
Colonoscopy reports[g]	–	–	–	–	–	–	–	–	–	–	✓

[a]FIT: fecal immunochemical test.

[b]Duration depends on how long insurants are sending back the FITs.

[c]Arm A: invitation + FIT; arm B: invitation + order; and arm C: information only (ie, control).

[d]Not applicable.

[e]Subrandomization in arms A and B: 50% receive reminder.

[f]Aggregated for all arms and additionally pseudonymized for arms A and B where consent was given.

[g]Anonymized for all arms where consent was given.

Inclusion and Exclusion Criteria

Inclusion criteria for receiving a letter are as follows: (1) aged 50-54 years, (2) enrolled in the AOK Family Doctor Program (HausarztProgramm) of AOK BW (Hausarztzentrierte Versorgung [HZV] contract: § 73b SGB V), and (3) did not receive an invitation in previous rounds of AOK BW model projects in previous years. Exclusion criteria for receiving a letter are as follows: (1) had an insurance-recorded colonoscopy within the past 5 years, (2) had a previous cancer diagnosis, (3) being deceased, and (4) being dependent on care.

Participants with a positive FIT result detected with the provided test kit are contacted by the DKFZ during the second part of the study to follow up on usage and outcome of colonoscopies.

Part 1: Personal Invitation for Colorectal Cancer Screening

The intervention is a personal invitation letter to perform a commercially available and validated FIT for quantitative detection of human hemoglobin (Hb); the test used is the OC-Sensor FIT (Eiken Chemical, Tokyo, Japan). Study information is being embedded in a mandatory, insurance-related information letter about the offer to undergo colonoscopies from

age 50 years, within the AOK Specialist Program. Suitable insured persons are being randomized into three arms—A, B, and C (ie, control)—by AOK BW (randomization was performed using a random number generator: structured query language [SQL] statement, database management system [DBMS]_RANDOM). Letters are being sent out biweekly in six tranches—one-sixth of each arm per tranche. A second randomization is dividing arms A and B into two respective subgroups—arm A: A1 and A2; arm B: B1 and B2. Subgroups A1 and B1 get a reminder letter after 4 weeks. No additional FIT kit is attached to the reminder letter and it states that the reminder is invalid if the FIT offer was already taken or ordered.

Within intervention arm A, insured persons receive a personal invitation letter, which describes CRC screening methods and includes a free FIT in a prepacked test kit. This kit contains one FIT, written and graphical instructions, two stool sample collectors, a record sheet for date of sample taking and year of birth, and a return envelope for mailing the test to the study center at DKFZ. Furthermore, the study information sheet, with optional email or phone consultation by the DKFZ, and the consent form are included.

Within intervention arm B, insured persons are also sent a personal invitation letter describing CRC screening methods, with the additional offer to order a FIT for free from the DKFZ via reply mail, fax, email, or online form. The study information sheet is added, but the consent form is only sent with the prepacked test kit upon FIT request.

Within control arm C, insured persons receive the insurance-related information letter mentioning a colonoscopy but do not receive the FIT itself nor the offer to order a FIT from the DKFZ. They may, however, make use of the FIT in routine practice: aggregated data on routine use is being provided by AOK BW.

Insurants randomized to arms A and B, who decide to take the FIT offer, are asked to sign the informed consent form. By signing this, they agree to the FIT analysis; to receive a written test result, which will also be sent to the family physician if not contradicted; to data storage and evaluation (ie, FIT result, age, gender, and first three digits of postal code); and to be recontacted during follow-up in case of a positive result. Optionally, they can allow the DKFZ to receive pseudonymized claims data from CRC screening participation in routine practice within 1 year after the initial invitation.

As proposed by the FIT manufacturer, participants are being instructed to spread the tip of the test stick over the freshly passed whole feces at various points until the tip’s grooves were filled with feces, and then to reinsert the probe into the device. The serrated probe that is attached to the device cap collects 10 mg feces into 2 mL of buffer. The FIT and the informed consent sheet are mailed to the DKFZ in a postpaid envelope. Upon arrival at the DKFZ, the FITs are stored in a fridge and separated from person-identifying information (ie, pseudonymized). The FITs are then sent by cooled transport to an external certified laboratory: Labor Limbach, Heidelberg, Germany, DIN EN ISO 15189 accredited. Trained laboratory personnel blinded to the randomization arm perform the analysis in a fully automated manner using the original OC-Sensor Pledia FIT device. Samples are disposed of after the analysis. As recommended by the manufacturer, the cutoff for a test result to be positive is 50 ng Hb/mL buffer (equal to 10 µg Hb/g feces). Results above the upper analytical limit of 1000 ng Hb/mL buffer are not diluted and not retested. Collection, arrival, and analysis dates of fecal samples are recorded. The DKFZ receives the laboratory reports, each checked and signed by a certified medical doctor. Additionally, the quantitative test results of the FITs are transferred electronically to the DKFZ, using DocNet plus, version 4 (DocNet Systems GmbH), where they are integrated into the study database.

Missing informed consent forms are being requested before any test analysis starts. Participants with failed tests receive a new test kit (eg, due to failed sampling or if the difference between the day of sample taking and arrival at the DKFZ is more than 7 days without cooling, followed by a negative test result).

The DKFZ sends a notification about the qualitative test result (ie, positive or negative, based on the manufacturer's recommended threshold) to the participants in an understandable manner. With a positive result, a consultation with the general practitioner is recommended to consider a colonoscopy for further examination. A copy of the laboratory report is sent to the family physician, unless participants did not wish so.

Part 2: Follow-Up of the Use of Screening Methods

A total of 6 months after a positive test result, the DKFZ contacts the respective participants in arms A and B to request permission for obtaining the reports of any subsequently conducted colonoscopies. Separate information sheets and consent forms are being used and the name of the treating gastroenterologist is requested. Relevant information from colonoscopy reports are being extracted and entered into an electronic database by two independent, trained data extractors and checked for inconsistencies.

Claims data of screening colonoscopies and FITs conducted within 1 year after the initial invitation or information letter is being provided by AOK BW in an aggregated manner per study arm (ie, A, B, and C) and gender. Individual pseudonymized claims data regarding the usage of FITs and colonoscopies is available for participants in the intervention arms upon specific informed consent.

In addition, results of screening colonoscopies in the age group of 50-54 years is being documented by the gastroenterologists in standardized survey forms, which are mandatory for billing. The billing company *MEDIVERBUND AG* captures the data electronically, anonymizes it, and sends it to the DKFZ for the analysis. Assignment to the initial study arms—A, B, or C—is possible only after consent.

Sample Size Calculation

The study population can be drawn in a very efficient manner from AOK BW. Approximately 17,532 insured persons met the inclusion criteria and received an initial letter: 5844 persons per arm—A, B, and C. A reminder is being sent out to 50% of the people in arms A and B—2922 persons per subgroups A1 and B1—after 4 weeks.

The following expectations are based on experiences from a previous model project [8], taking into account changes in this study's conditions. The expected usage of a FIT for both genders combined is about 10% for the people in the control group (ie, FIT use in routine practice). In intervention arm A (ie, invitation with FIT), a 3.5-fold increase to 35% (ie, about 30% after the first letter plus about 5% after the reminder) is expected. In intervention arm B (ie, invitation with offer to order a FIT), a 1.75-fold increase to about 17.5% is expected (ie, about 15% after the first letter plus about 2.5% after the reminder). The sample size—approximately 5844 per arm—allows a high-precision estimation of the expected relative increase in FIT usage; the expected 95% CIs for relative FIT usage are 3.22-3.81 and 1.59-1.92 in intervention arms A and B, respectively, compared to the control arm. The statistical power to detect an effect within the expected range is close to 100% for both types of intervention: SAS, version 9.4 (SAS Institute) POWER procedure, two-sided Pearson chi-square test with significance level alpha=.017. For the comparison of the three arms—A, B, and C—we adjusted for multiple testing according to the Bonferroni-Holm method. For the most stringent adjustment with alpha=.017 (ie, baseline alpha=.05 with three tests), there is still excellent power.

In the randomized substudy, the increase in FIT usage due to a reminder letter in comparison to the one-time invitation will be analyzed. The expected FIT usage in arm A is approximately 40% with the reminder compared to 30% without the reminder; the expected FIT usage in arm B is approximately 20% with the reminder compared to 15% without the reminder. The included sample size of the randomized substudy (ie, 2922 per arm) allows for a precise estimation of the impact of sending a reminder. The expected relative FIT use (95% CI) in subgroups A1 and B1 (ie, received a reminder) compared to the respective subcontrol groups A2 and B2 (ie, one-time invitation only) would be 1.33 (1.24-1.43) and 1.33 (1.19-1.49), respectively. The statistical power to detect an effect of the expected order of magnitude is close to 100%: SAS, version 9.4 (SAS Institute) POWER procedure, two-sided Pearson chi-square test with significance level alpha=.025. For the two comparisons (ie, A1 vs A2 and B1 vs B2), multiple testing was adjusted for, according to the Bonferroni-Holm method. For the most stringent adjustment with alpha=.025 (ie, baseline alpha=.05 with two tests) there is still excellent power.

Statistical Analysis

The statistical analysis of this confirmatory study primarily involves comparing the use of FITs in the intervention arms and the control arm within 1 year after the initial letter. In addition, within the intervention arms, the effect of a one-time invitation will be compared to an invitation with a reminder letter.

Comparisons will be done by two-sided chi-square tests for differences in the participation rates. Secondary outcomes will be addressed by descriptive and exploratory analyses. Data for the primary outcome analysis are expected to be complete, as complete information on FITs conducted in routine practice is obtained through insurance claims data (ie, billing codes for laboratory analysis of the FIT), and complete information on FITs conducted through the special offers is directly available from the study center.

Results

In total, 17,532 invitation letters and 5825 reminder letters were sent out; the study was launched in September 2017. Data collection and workup were completed in fall 2019.

Discussion

We initiated this three-armed randomized controlled trial in order to evaluate, with the highest possible evidence [11,12], the effect of low-threshold invitation schemes on the usage of CRC screening—with focus on FITs—in the 50-54-year-old population.

CRC remains one of the most frequent causes of cancer and reasons for cancer-related death in Germany [13] and worldwide [14]. Participation rates of CRC screening need to be increased; various approaches have recently been investigated in different countries in a rising number of population-based [7] and randomized trials [8,15-17]. Among those are results of a nationwide, FIT-based screening program in the Netherlands [7], as well as from a previously conducted model project in Germany [8]; these studies have consistently shown that the usage of a stool test increases after the target population receives personal invitations including the test. Despite that, the nationwide organized invitation procedure, which was introduced in Germany in 2019 to improve CRC screening, only comprises a personal invitation sent out by health insurance plans, with an enclosed gender-specific information brochure but no direct provision of, or low-threshold access to, a FIT [3]. Moreover, first trends indicate that the use of a stool test has further declined since the change from gFOBT to the more sensitive and widely recommended FITs in 2017 [18-20].

Usage of the available screening methods for CRC reduces the incidence and the mortality of CRC by removal of precursors and detection of cancer at an early stage [1,21] and, thereby, reduces the costs associated with CRC-related therapies. Furthermore, colonoscopies following a positive FIT result might be more effective compared to colonoscopies without a previous FIT, due to a higher chance of detecting relevant findings, such as advanced adenomas. Thus, an invitation procedure with low-threshold FIT provision could not only increase the usage of the test, but also improve the effectiveness of subsequent colonoscopies, thereby avoiding potential adverse outcomes and costs arising in the case of later detection of CRC.

The concept of this study allows for monitoring and investigating the usage rate of FITs following a once-only CRC screening information letter in routine practice and after invitation with provision of the FIT; whether a colonoscopy was conducted after the FIT, no matter if there was a positive or negative FIT outcome; as well as the outcome of such a colonoscopy. This might lead to a better understanding of the screening-related and possibly gender-specific actions taken as well as the needs in the targeted population. Findings from colonoscopy reports might further help to point out the relevance

for screening, due to the prevalence of advanced adenomas in this age group.

All invited persons are between 50 and 54 years of age, representing the age group for whom the FIT is covered in the national screening program. Nevertheless, selection bias due to the focus of our study on AOK-insured persons enrolled in the AOK Family Doctor Program cannot be ruled out, which may limit external validity of the study.

To our knowledge, this is the first study that examines a low-threshold order option of the FIT, in addition to direct provision of a FIT, included in the invitation letter. The concept of personal invitations being sent out by a health insurance plan is currently implemented in Germany [3]; however, it includes neither direct provision of a FIT with the invitation letter nor a low-threshold order option. The results of this randomized controlled trial will, therefore, provide important empirical evidence for potential further enhancement of CRC screening in routine practice in the German health care system and beyond.

Acknowledgments

We gratefully acknowledge the excellent cooperation with AOK BW in the design and implementation of the study, with special thanks to Stefan Meny. We would further like to thank Labor Limbach for the reliable test analyses as well as the laboratory and documentary staff of our division for their great work and support. The study is funded by AOK BW. Since the study aims to explore settings that might be involved in routine practice, AOK BW was involved in the study design and data collection (ie, randomization, initial contact with insurants, and data from routine practice). The funding source plays neither a role in data analysis and interpretation nor in final decisions related to publications.

Authors' Contributions

LFG contributed to the design, preparation, and conduct of the study and drafted the manuscript. MH and LL contributed to the design of the study and the revision of the manuscript. HB designed, led, and supervised the study and contributed to the revision of the manuscript. All authors approved the final, submitted version of the manuscript.

Conflicts of Interest

None declared.

References

1. Shaukat A, Mongin SJ, Geisser MS, Lederle FA, Bond JH, Mandel JS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013 Sep 19;369(12):1106-1114. [doi: 10.1056/NEJMoa1300720] [Medline: 24047060]
2. European Colorectal Cancer Screening Guidelines Working Group, von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis: Overview and introduction to the full supplement publication. Endoscopy 2013;45(1):51-59 [FREE Full text] [doi: 10.1055/s-0032-1325997] [Medline: 23212726]
3. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses für organisierte Krebsfrüherkennungsprogramme. 2019 Jul. URL: https://www.g-ba.de/downloads/62-492-1844/oKFE-RL-2018-11-22-iK-2019-07-01.pdf [accessed 2019-07-05]
4. Lopes G, Stern MC, Temin S, Sharara AI, Cervantes A, Costas-Chavarri A, et al. Early detection for colorectal cancer: ASCO resource-stratified guideline. J Glob Oncol 2019 Feb;5:1-22 [FREE Full text] [doi: 10.1200/JGO.18.00213] [Medline: 30802159]
5. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: A global overview of existing programmes. Gut 2015 Oct;64(10):1637-1649. [doi: 10.1136/gutjnl-2014-309086] [Medline: 26041752]
6. Senore C, Basu P, Anttila A, Ponti A, Tomatis M, Vale DB, et al. Performance of colorectal cancer screening in the European Union Member States: Data from the second European screening report. Gut 2019 Jul;68(7):1232-1244. [doi: 10.1136/gutjnl-2018-317293] [Medline: 30530530]
7. Kapidzic A, Grobbee EJ, Hol L, van Roon AH, van Vuuren AJ, Spijker W, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. Am J Gastroenterol 2014 Aug;109(8):1257-1264. [doi: 10.1038/ajg.2014.168] [Medline: 24980879]
8. Hoffmeister M, Holleczek B, Zwink N, Stock C, Stegmaier C, Brenner H. Screening for bowel cancer: Increasing participation via personal invitation. Dtsch Arztebl Int 2017 Feb 10;114(6):87-93 [FREE Full text] [doi: 10.3238/arztebl.2017.0087] [Medline: 28266301]
9. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. BMJ 2013 Jan 08;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
10. Fraser CG, Allison JE, Young GP, Halloran SP, Seaman HE. Improving the reporting of evaluations of faecal immunochemical tests for haemoglobin: The FITTER standard and checklist. Eur J Cancer Prev 2015 Jan;24(1):24-26. [doi: 10.1097/CEJ.0000000000000016] [Medline: 24584197]

11. Kabisch M, Ruckes C, Seibert-Grafe M, Blettner M. Randomized controlled trials: Part 17 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2011 Sep;108(39):663-668 [FREE Full text] [doi: 10.3238/arztebl.2011.0663] [Medline: 22013494]
12. Misra S. Randomized double blind placebo control studies, the "Gold Standard" in intervention based studies. Indian J Sex Transm Dis AIDS 2012 Jul;33(2):131-134 [FREE Full text] [doi: 10.4103/0253-7184.102130] [Medline: 23188942]
13. The Cancer Register Data Center, The Society of the Epidemiological Cancer Register in Germany. Krebs in Deutschland für 2013/2014. Berlin, Germany: Robert Koch-Institut; 2017. URL: https://tinyurl.com/t82apvl [accessed 2020-01-22]
14. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Nov;68(6):394-424 [FREE Full text] [doi: 10.3322/caac.21492] [Medline: 30207593]
15. Hirai K, Ishikawa Y, Fukuyoshi J, Yonekura A, Harada K, Shibuya D, et al. Tailored message interventions versus typical messages for increasing participation in colorectal cancer screening among a non-adherent population: A randomized controlled trial. BMC Public Health 2016 May 24;16:431 [FREE Full text] [doi: 10.1186/s12889-016-3069-y] [Medline: 27220976]
16. Denis B, Broc G, Sauleau EA, Gendre I, Gana K, Perrin P. Tailored telephone counselling to increase participation of underusers in a population-based colorectal cancer-screening programme with faecal occult blood test: A randomized controlled trial. Rev Epidemiol Sante Publique 2017 Feb;65(1):17-28. [doi: 10.1016/j.respe.2016.06.336] [Medline: 28089385]
17. Mehta SJ, Pepe RS, Gabler NB, Kanneganti M, Reitz C, Saia C, et al. Effect of financial incentives on patient use of mailed colorectal cancer screening tests: A randomized clinical trial. JAMA Netw Open 2019 Mar 01;2(3):e191156 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.1156] [Medline: 30901053]
18. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungsrichtlinie: Bewertung eines iFOBT-basierten Darmkrebsscreenings im Vergleich zu einem gFOBT-basierten Darmkrebsscreening. 2016 Apr. URL: https://www.g-ba.de/downloads/40-268-3744/2016-04-21_KFE-RL_Bewertung-iFOBT_TrG.pdf [accessed 2019-06-24]
19. Altenhofen L. Projekt Wissenschaftliche Begleitung von Früherkennungs-Koloskopien in Deutschland Berichtszeitraum 2014. Köln (Cologne), Germany: Zentralinstitut für die kassenärztliche Versorgung in Deutschland; 2016 Jan. URL: https://www.zi.de/fileadmin/user_upload/Jahresbericht_2014_Darmkrebs_Frueherkennung.pdf [accessed 2019-09-25]
20. Gemeinsamer Bundesausschuss. Der iFOBT im Darmkrebs-Screening: Ergebnisse der medizinischen Laboratorien für das Jahr 2018. 2019. URL: https://www.g-ba.de/downloads/17-98-4777/2019-03-25_G-BA_iFOBT_Quartalsbericht_2018_.pdf [accessed 2019-06-17]
21. Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: Systematic review and meta-analysis of randomised controlled trials and observational studies. BMJ 2014 Apr 09;348:g2467 [FREE Full text] [doi: 10.1136/bmj.g2467] [Medline: 24922745]

Abbreviations

AOK BW: AOK Baden-Wuerttemberg
CRC: colorectal cancer
DBMS: database management system
DKFZ: German Cancer Research Center
DRKS: German Clinical Trials Register
FIT: fecal immunochemical test
FITTER: Fecal Immunochemical TesTs for Hemoglobin Evaluation Reporting
FOBT: fecal occult blood test
gFOBT: guaiac-based fecal occult blood test
Hb: hemoglobin
HZV: Hausarztzentrierte Versorgung
iFOBT: immunochemical fecal occult blood test
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SQL: structured query language

Edited by G Eysenbach; submitted 26.09.19; peer-reviewed by M Dreier, R Kerrison, A Benslimane; comments to author 06.11.19; revised version received 20.11.19; accepted 26.11.19; published 03.04.20.

Please cite as:
Gruner LF, Hoffmeister M, Ludwig L, Brenner H
Effect of Various Invitation Schemes on the Use of Fecal Immunochemical Tests for Colorectal Cancer Screening: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e16413
URL: https://www.researchprotocols.org/2020/4/e16413
doi:10.2196/16413
PMID:32242518

Protocol

Evaluation of the Safety and Efficacy of Avacopan, a C5a Receptor Inhibitor, in Patients With Antineutrophil Cytoplasmic Antibody–Associated Vasculitis Treated Concomitantly With Rituximab or Cyclophosphamide/Azathioprine: Protocol for a Randomized, Double-Blind, Active-Controlled, Phase 3 Trial

Peter A Merkel[1,2], MD, MPH; David R Jayne[3], MD; Chao Wang[4], PhD; Jan Hillson[5], MD; Pirow Bekker[5], MD

[1]Division of Rheumatology, Department of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Division of Clinical Epidemiology, Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA, United States

[3]Department of Medicine, University of Cambridge, Cambridge, United Kingdom

[4]Biostatistics, Pharma Data Associates, LLC, Piscataway, NJ, United States

[5]Research and Development, ChemoCentryx, Inc, Mountain View, CA, United States

Corresponding Author:
Peter A Merkel, MD, MPH
Division of Rheumatology
Department of Medicine
University of Pennsylvania
5th Floor, White Building
3400 Spruce Street
Philadelphia, PA, 19104
United States
Phone: 1 215 614 4401
Email: pmerkel@upenn.edu

Abstract

Background: Antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis is a serious, often life-threatening disease. In new-onset disease or a relapse, the standard treatment is immunosuppressive therapy with glucocorticoids; these therapies are associated with substantial short- and long-term toxicity. Complement component 5a (C5a) binding to C5a receptor (C5aR) may play a central role in the pathogenesis of ANCA-associated vasculitis. Avacopan is a novel, orally bioavailable, and highly selective antagonist of human C5aR. Avacopan does not interfere with the production of C5b or the membrane attack complex (ie, terminal complement complex) and does not block C5a binding to a second receptor, C5L2 (also called C5aR2), shown to be protective in antimyeloperoxidase glomerulonephritis. This trial will evaluate if avacopan replaces the need for chronic glucocorticoids in the treatment of ANCA-associated vasculitis.

Objective: The aim of this study is to determine the proportions of patients in remission at week 26 and with sustained remission at week 52, defined as Birmingham Vasculitis Activity Score=0, and not taking glucocorticoids within the 4 weeks before week 26 and week 52, respectively.

Methods: The Avacopan Development in Vasculitis to Obtain Corticosteroid elimination and Therapeutic Efficacy study is a randomized, double-blind, active-comparator (prednisone), 2-arm study evaluating the safety and efficacy of avacopan versus prednisone, administered in combination with other immunosuppressive therapy. Eligible subjects will have active disease requiring induction of remission. Subjects are stratified based on the type of immunosuppressive therapy, ANCA subtype, and new or relapsing disease. Target sample size is 300 patients, enrolled at over 200 sites globally. All authors and local ethics committees approved the study design. All patients will provide informed consent.

Results: Enrollment of patients was completed in Q4 2018. Topline results are anticipated to be published by Q3 2020.

Conclusions: Results will be released irrespective of whether the findings are positive or negative.

Trial Registration: ClinicalTrials.gov NCT02994927; https://clinicaltrials.gov/ct2/show/NCT02994927

International Registered Report Identifier (IRRID): DERR1-10.2196/16664

(JMIR Res Protoc 2020;9(4):e16664) doi:10.2196/16664

KEYWORDS

ANCA-associated vasculitis; avacopan; C5a receptor; ADVOCATE

Introduction

Background

Antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis is a serious, often life-threatening disease that includes three related forms of small-vessel vasculitis: granulomatosis with polyangiitis (GPA), microscopic polyangiitis (MPA), and eosinophilic granulomatosis with polyangiitis (Churg-Strauss). GPA and MPA are thought to be triggered by the production of circulating autoantibodies against the neutrophil-expressed antigens myeloperoxidase (MPO) or proteinase 3 (PR3). The disease may present with a wide spectrum of clinical manifestations of varying degrees of severity including but not limited to skin lesions, sinonasal, pulmonary inflammation or hemorrhage, and glomerulonephritis.

The proinflammatory complement C5a ligand and its receptor C5aR (also referred to as CD88) appear to play a central role in the pathogenesis of ANCA-associated vasculitis [1-4]. When stimulated by inflammatory cytokines, activation of the terminal C5a/C5aR axis generates an autoamplification loop that drives the acute necrotizing vasculitic process from primed neutrophils resulting from their interaction with ANCAs. C5a acting on C5aR is a potent neutrophil chemoattractant and agonist [5], which increases neutrophil adhesion, induces neutrophil degranulation, and produces reactive oxygen intermediates. Activation of C5aR also decreases neutrophils' deformability, slowing their ability to traverse small blood vessels particularly in the presence of ANCA [6]. Moreover, C5a activates vascular endothelial cells through C5aR, promoting their retraction and increased vascular permeability [7,8].

The current standard treatments for induction of remission for GPA and MPA rely on substantial immunosuppression and consist primarily of glucocorticoids combined with either cyclophosphamide (intravenous or oral) followed by oral azathioprine, or rituximab [9-11]. Despite these available therapies, patients with ANCA-associated vasculitis have a 9-fold increased risk of mortality in the first year compared with healthy controls, attributed to infection, active vasculitis, and the effects of renal insufficiency [12,13]. More than 50% of the 1-year mortality is attributed to treatment-related adverse events rather than active vasculitis [12,13]. Furthermore, substantial accumulated permanent damage can occur when the disease is insufficiently controlled; for example, 15% to 38% of patients develop end-stage renal disease within 5 years [14-19].

Although glucocorticoids have a relatively rapid onset of action, their use is associated with an overall negative impact on the patient's health and health-related quality of life (HRQoL), attributable to known acute and chronic effects of these medications [20]. In the treatment of ANCA-associated vasculitis, these include increased risks of infection [21-23], diabetes mellitus, fractures, gastrointestinal bleeding, hypertension, cataracts, and progressive organ damage [24]. Weight gain, sleep disturbance, lipodystrophy, and neuropsychiatric disturbances (including irritability, anxiety, depression, and hyperactivity) have also been reported [25,26]. Moreover, an analysis of data from patients enrolled in several trials conducted by the European Vasculitis Study Group showed that the level of damage associated with the use of glucocorticoids in ANCA-associated vasculitis increased with length of exposure [24]. In relapsing disease, the repeated use of glucocorticoids leads to even further damage. Collectively, the data indicate a need for effective glucocorticoid-sparing treatments that can help avoid complications associated with exposure to glucocorticoids and limit organ damage through more rapid and sustained disease control.

Avacopan (previously known as CCX168) is a novel, orally bioavailable, highly selective human C5aR antagonist that lacks any other known pharmacology [27]. Importantly, avacopan does not inhibit the interaction of C5a with the related receptor C5L2 (also referred to as C5aR2), thought to have anti-inflammatory properties [27] and also shown to impart protective effects in a mouse model of anti-MPO-induced glomerulonephritis [28]. As a specific C5aR antagonist, avacopan does not interfere with the formation of the terminal complement complex or membrane attack complex C5b-9, which is necessary for clearance of pathogenic encapsulated bacteria such as *Neisseria meningitidis*.

Prior Work

Along with a favorable safety profile in a phase 1 dose escalation clinical trial in healthy volunteers, 30 mg of avacopan administered twice daily was shown to produce plasma concentrations that provided near complete inactivation of C5aR on blood neutrophils throughout the day, including ≥90% C5aR blockade at trough levels of the drug [27]. This exposure was sufficient to achieve optimal blockade of C5a-induced CD11b upregulation, as well as C5a-induced degranulation, reactive oxygen intermediate production, and C5a-directed migration of neutrophils ex vivo, forming the basis for the dose selected for avacopan in subsequent clinical studies of ANCA-associated vasculitis.

In a phase 2 clinical trial (NCT01363388) conducted in ANCA-associated vasculitis, 67 patients received one of the three treatments: avacopan with no prednisone, avacopan with low-dose (20 mg/day) prednisone (with taper), or placebo control with standard dose (60 mg/day) prednisone (with taper); treatment was administered along with either cyclophosphamide followed by azathioprine or rituximab. Avacopan was shown to replace glucocorticoids without compromising efficacy. Clinical response at week 12, defined as a 50% reduction in Birmingham Vasculitis Activity Score (BVAS) from baseline,

was observed in 70% of the standard treatment group (placebo), 86% of the avacopan plus low-dose prednisone group (*P*=.002 for noninferiority, compared with standard care), 81% of the avacopan without prednisone group (*P*=.01 for noninferiority, compared with standard care), with a trend for faster control of disease in the avacopan groups relative to control [29]. Patients in the avacopan groups also had a rapid improvement in albuminuria (an indicator of preserved renal function), whereas the estimated glomerular filtration rate and hematuria improved in all treatment groups. Fewer glucocorticoid-associated adverse effects (primarily driven by a lower incidence of new-onset or worsening diabetes mellitus and psychiatric disorders) along with improvements in HRQoL (Medical Outcomes Study Short-Form 36 [SF-36] version 2, and EuroQOL [EQ-5D-5L]) were also associated with avacopan.

A second phase 2 trial of 12-week duration in 42 patients with ANCA-associated vasculitis (NCT02222155) was designed to examine the safety of avacopan when combined with standard of care therapy including glucocorticoids. Comparison of the avacopan versus placebo groups revealed that avacopan therapy when added to standard of care added no new safety signals beyond those associated with standard of care [30].

Goals and Objectives of This Study

These phase 2 studies suggest that avacopan is well tolerated and that specific inhibition of the C5a/C5aR interaction on inflammatory cells may improve outcomes in ANCA-associated vasculitis, while reducing glucocorticoid exposure and their associated side effects observed with the standard of care. On the basis of these results, the underlying hypothesis for the phase 3 clinical trial protocol presented here (NCT02994927) is that avacopan is at least as effective as glucocorticoids in rapidly inducing and sustaining remission of signs and symptoms of ANCA-associated vasculitis when combined with either cyclophosphamide followed by azathioprine or with rituximab. As a selective anti-inflammatory agent, avacopan may also have safety advantages compared with glucocorticoids.

The primary objective of the trial is to evaluate the efficacy of avacopan to induce and sustain remission in patients with ANCA-associated vasculitis, when used in combination with immunosuppressive therapy, rituximab, or cyclophosphamide followed by azathioprine (or mycophenolate mofetil), but in the absence of months of treatment with glucocorticoids. As many patients in the trial are expected to have renal involvement due their vasculitis and methotrexate would be contraindicated in such patients, methotrexate was not an option for use following cyclophosphamide; this stipulation also aimed to reduce variation in the treatment groups. Disease remission is defined as achieving a BVAS score [31,32] of 0 (no evidence of active vasculitis) and no use of glucocorticoids for the treatment of ANCA-associated vasculitis within 4 weeks before week 26. Sustained remission is defined as remission at week 26 without relapse to week 52 (BVAS=0) and no use of glucocorticoids for ANCA-associated vasculitis within 4 weeks before week 52.

Secondary objectives include evaluation of the effect of treatment with avacopan versus standard of care on overall safety, glucocorticoid-related toxicity, rapidity of response, changes in HRQoL, changes in renal disease, and cumulative organ damage (vasculitis damage index).

Methods

Study Design

Enrollment of 300 patients is planned across more than 200 sites in 20 countries on 4 continents. Patients who drop out following randomization are not replaced. Briefly, after an explanation of the study and obtaining written informed consent, potentially eligible patients are screened against the inclusion and exclusion criteria (Textboxes 1 and 2) to confirm eligibility including review of their medical history.

Textbox 1. Major inclusion criteria for the randomization study.

- Age ≥12 years
- Newly diagnosed or relapsing antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis with granulomatosis with polyangiitis or microscopic polyangiitis
- Positive for antiproteinase 3 or antimyeloperoxidase ANCA
- Active disease, as assessed by ≥1 major item and ≥3 nonmajor items, or ≥2 renal items on Birmingham Vasculitis Activity Score (Multimedia Appendix 1) [31,32]
- Estimated glomerular filtration rate ≥15 mL/min/1.73 m^2

Patients are required to fast for at least 9 hours before the first dose of study drug to allow for baseline low-density lipoprotein–cholesterol measurements. Upon arriving at the study site, patients undergo baseline assessments including sample collections to test for autoimmune serologies, hematologic parameters, a comprehensive chemistry panel, urinalysis, and additional blood and urine sampling for research purposes; HRQoL questionnaires (SF-36 and EQ-5D-5L) and glucocorticoid toxicity index [33] are completed.

Textbox 2. Major exclusion criteria for the randomization study.

- Any of the following conditions: alveolar hemorrhage requiring invasive pulmonary ventilation support, other multisystem autoimmune disease (including eosinophilic granulomatosis with polyangiitis, lupus, IgA vasculitis (Henoch-Schönlein), rheumatoid vasculitis, Sjögren's syndrome, antiglomerular basement membrane disease, or cryoglobulinemic vasculitis)
- Requires dialysis or plasma exchange within 12 weeks before screening
- History of kidney transplant
- Immunosuppressive therapies: received cyclophosphamide ≤12 weeks before screening
- On azathioprine, methotrexate, or mycophenolate mofetil at screening and unwilling to discontinue use and switch to cyclophosphamide or rituximab on day 1
- Received intravenous glucocorticoids, >3000 mg methylprednisolone equivalent, within 4 weeks before screening or oral glucocorticoid >10 mg prednisone-equivalent within 6 weeks continuously before screening
- Received rituximab or other B-cell antibody ≤52 weeks before screening, or ≤26 weeks before screening provided B cell reconstitution has occurred, that is, CD19 count >0.01 × 109/L)
- Received antitumor necrosis factor treatment <12 weeks before screening
- For patients scheduled to receive cyclophosphamide: urinary outflow obstruction, active infection, or platelet count <50,000/μL before start of dosing
- Previous receipt of avacopan
- History of cancer within last 5 years, with the exception of excised basal cell or squamous cell carcinoma of the skin, or carcinoma in situ such as cervical or breast carcinoma in situ that has been excised or resected completely and is without evidence of local recurrence or metastasis
- Evidence of hepatic disease (transaminases, alkaline phosphatase >3 times upper limit of normal)
- Known active infection with tuberculosis, hepatitis B or C virus, or human immunodeficiency virus
- White blood cell count <3500/μL or neutrophil count <1500/μL, or lymphocyte count <500/μL at baseline

Stratification and Treatment Groups

Eligible patients with ANCA-associated vasculitis are stratified according to 3 parameters: (1) baseline immunosuppressive therapy (oral cyclophosphamide, intravenous cyclophosphamide, or rituximab), (2) type of ANCA (anti-PR3 or anti-MPO), and (3) whether the patient is newly diagnosed or has relapsing disease (Figure 1). Following stratification, patients are randomized in a 1:1 ratio to 2 treatment groups receiving baseline immunosuppressive therapy plus either (A) blinded avacopan 30 mg twice daily plus placebo that matches prednisone or (B) blinded placebo that matches avacopan plus prednisone starting at 60 mg/day (or adjusted for weight). In the (B) group (standard of care), study-supplied prednisone is tapered to zero by day 140 (Table 1). Additional (nonstudy supplied) glucocorticoids are allowed before enrollment and during the first 4 weeks if required for initial control of disease as outlined later (*Additional Glucocorticoids*). In addition, the protocol allows for the use of low-dose glucocorticoids (up to 10 mg/day prednisone or equivalent) to treat adrenal insufficiency or to treat worsening or relapsing disease. A minimization algorithm is used to dynamically assign patients to a treatment group, which seeks to balance treatment groups with respect to each stratification factor. In adolescents (aged 12-17 years), the initial dose of avacopan can be adjusted downward (to 10 or 20 mg twice daily) depending on body weight with potential refinement based on plasma levels of avacopan. Any dose changes are determined by an unblinded reviewer, not otherwise associated with study, following the first dose of study drug; to maintain the blind, some patients receiving placebo will also have their dose modified.

Figure 1. Avacopan Development in Vasculitis to Obtain Corticosteroid Elimination and Therapeutic Efficacy trial design. AZA: azathioprine; BVAS: Birmingham Vasculitis Activity Score; CYC: cyclophosphamide; GC: glucocorticoids; IV: intravenous; MPO: myeloperoxidase; PR3: proteinase 3; RTX: rituximab.

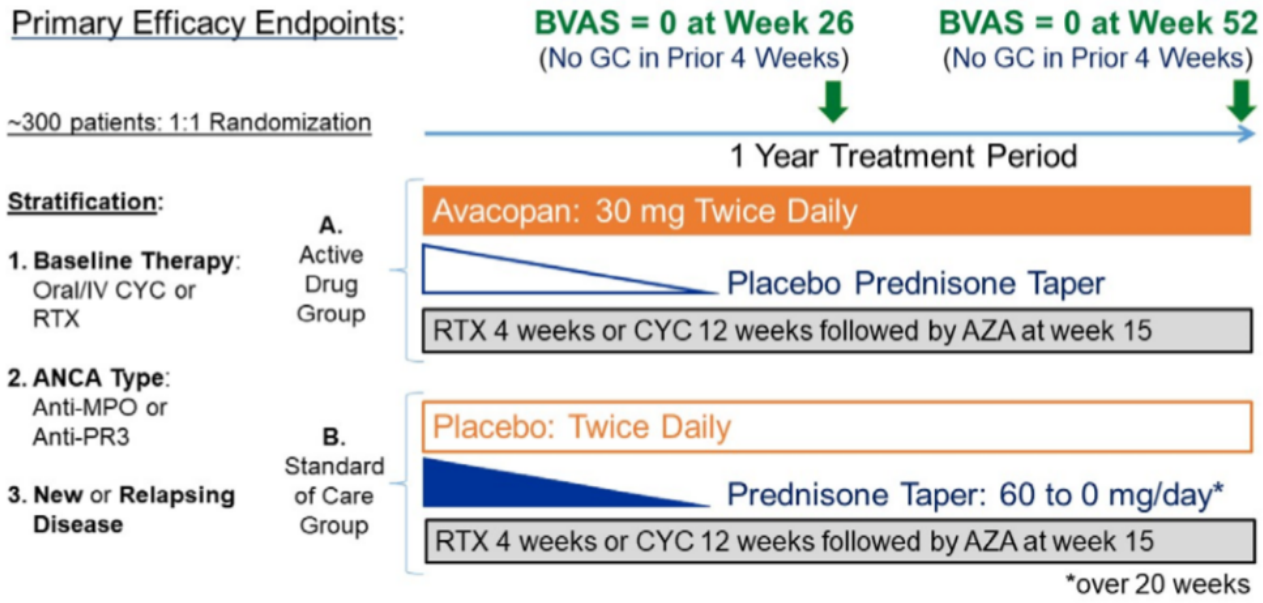


Table 1. Glucocorticoid and matching placebo tapering schedule.

Study day	Group A: avacopan (prednisone dose, mg per day)	Group B: prednisone (control; prednisone dose, mg per day)			
	All patients	Adults		Adolescents (aged 12-17 years)	
		≥55 kg	<55 kg	>37 kg	≤37 kg
1-7	0	60	45	45	30
8-14	0	45	45	45	30
15-21	0	30	30	30	30
22-42	0	25	25	25	25
43-56	0	20	20	20	20
57-70	0	15	15	15	15
71-98	0	10	10	10	10
99-140	0	5	5	5	5
≥141	0	0	0	0	0

Immunosuppressive Therapy

All patients in both treatment groups also receive one of the following 3 investigator-selected baseline immunosuppressive therapy (standard-of-care) regimens:

1. Intravenous rituximab: 375 mg/m^2 weekly × 4 infusions.
2. Intravenous cyclophosphamide: 15 mg/kg up to 1.2 g every 2 to 3 weeks for 13 weeks, and then starting on week 15, oral azathioprine 1 mg/kg daily with titration up to 2 mg/kg daily (mycophenolate mofetil 2 g daily is allowed in place of azathioprine).
3. Oral cyclophosphamide: 2 mg/kg daily for 14 weeks followed by oral azathioprine or mycophenolate mofetil starting at week 15 (same dosing regimen as intravenous cyclophosphamide).

Dose reductions or adjustments in cyclophosphamide, azathioprine, and mycophenolate are allowed to conform to standard approaches to maximize safety of these medications.

Additional Glucocorticoids

Use of additional glucocorticoids (eg, nonstudy supplied glucocorticoids in excess of dose and schedule outlined in Table 1) is to be avoided during the study, except for cases of adrenal insufficiency. Patients currently on glucocorticoids may enroll provided they meet the following criteria: before and during the screening visit, the use of intravenous glucocorticoids does not exceed a cumulative dose equivalent of 3 g of methylprednisolone in the 4 weeks before screening or the use of oral glucocorticoids does not exceed 10 mg once daily (QD) of oral prednisone for up to 6 weeks before screening. During the screening period (≤14 days), patients will be allowed oral glucocorticoids (tapered to ≤20 mg QD prednisone equivalent by day 1). If a patient enrolls in the study while on oral

glucocorticoids, this dose must be tapered to 0 mg by the end of week 4.

Patients who experience worsening of disease during the study that involves a major item in the BVAS may be treated with intravenous glucocorticoids (typically 0.5-1 g methylprednisolone per day for 3 days) or oral glucocorticoids, tapered according to the patient's condition, or both. Worsening not involving a major item in the BVAS may be treated with a short burst (≤2 weeks) of oral glucocorticoids, at a maximum daily dose of 20 mg prednisone equivalent. The use of other medications, such as additional rituximab or cyclophosphamide, is discussed with the medical monitor before implementing. The use of plasma exchange is not permitted.

Patients experiencing a relapse or worsening of disease may continue study drug treatment and should continue in the study. In patients experiencing a relapse, the study-supplied prednisone/matching placebo will be temporarily halted during the course of glucocorticoids, and if the patient's condition stabilizes, the study-supplied prednisone/matching placebo may be restarted according to the original study visit schedule. At the discretion of the investigator, the avacopan/matching placebo may be continued during and following the treatment for the relapse.

Assessments, Endpoints, and Outcomes

BVAS assessments are performed at screening and weeks 4, 10, 16, 26, 39, 52, and 60. Vasculitis damage index assessments [34] are performed at screening and weeks 26, 52, and 60. HRQoL assessments are completed on day 1 and weeks 4, 10, 16, 26, 39, 52, and 60. Glucocorticoid toxicity index assessments are performed on day 1 and at weeks 13 and 26. If a patient consents, optional renal biopsies for histology are collected at baseline if not already on file, and at week 52 or at the time of treatment discontinuation. Physical examinations, vital sign assessments, and electrocardiogram measurements are performed throughout the study. Concomitant medication and adverse events are assessed at every study visit (Multimedia Appendix 1). Determination of whether a patient enters remission and sustains remission, as well as all relapses, is assessed by an independent blinded adjudication committee.

The key clinical endpoints and outcome measures are summarized in Textboxes 3 and 4. The overall efficacy hypothesis in this study is that avacopan, in combination with immunosuppressive therapy, will be at least as effective for treatment of ANCA-associated vasculitis when compared with prednisone with immunosuppressive therapy, and may reduce the overall burden of disease by achieving remission with less exposure to glucocorticoids, fewer glucocorticoid-associated adverse effects, and improved quality of life.

Textbox 3. Main clinical endpoints of this study.

- Proportion of patients in remission at week 26; defined as Birmingham Vasculitis Activity Score=0 and not taking glucocorticoids for antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis within 4 weeks before week 26
- Proportion of patients achieving sustained remission at week 52; defined as remission at week 26 and week 52, without relapse through week 52, and not taking glucocorticoids for ANCA-associated vasculitis within 4 weeks before week 52

Textbox 4. Key outcome measures of the study.

- Safety: adverse events, physical exam, vital signs, serum chemistry, hematology, urinalysis, and electrocardiogram
- Change in glucocorticoid-induced toxicity measured using the change from baseline in the glucocorticoid toxicity index version 2 cumulative worsening score and aggregate improvement score; cumulative use of glucocorticoids will also be analyzed. Rapidity of response based on early remission (Birmingham Vasculitis Activity Score [BVAS]=0) at week 4
- Change from baseline in health-related quality of life scores over 52 weeks based on the short form 36 version 2 and the Euro Quality of Life
- Proportion of patients and time to disease relapse after previously having achieved remission at week 26; relapse is defined as ≥1 major item in BVAS, ≥3 minor items in BVAS, or 1 or 2 minor items in BVAS at 2 consecutive visits
- Change in damage from baseline over 52 weeks, as measured by the Vasculitis Damage Index
- In patients with renal disease at baseline, change over 52 weeks in estimated glomerular filtration rate, urinary albumin: creatinine ratio, and urinary monocyte chemoattractant protein-1:creatinine ratio

Statistical Analysis

A minimization algorithm is used to assign patients dynamically to a treatment group, which seeks to balance treatment groups with respect to each stratification factor. Statistical analysis for categorical endpoints will be expressed as the number and percentage of unique patients for each category. For continuous variables, numbers, means, medians, ranges, SDs, and standard error of means will be calculated. Geometric means will be calculated for data that are not normally distributed. Results will be presented by treatment group and by stratum for each of the 3 stratification factors. Results will be presented for patients with and without renal involvement of disease at baseline.

To estimate sample size, a noninferiority margin of −20% was based on a thorough review and meta-analysis of all previous trials conducted in patients with ANCA-associated vasculitis as well as precedent. The proportion of patients with remission in the control group at week 26 was estimated to be 60%, which is based on a blended proportion of 64% and 53% observed in the rituximab and cyclophosphamide/azathioprine groups,

respectively, in the RAVE study [10]. On the basis of these assumptions and a 2-sided significance level of 0.05, a sample size of 150 patients per group provides >90% power for the noninferiority analysis. For the determination of the primary efficacy endpoints, the proportion of patients achieving disease remission at week 26 and sustained disease remission at week 52, the 2-sided 95% CIs for the difference in proportions (avacopan group minus prednisone group) will be estimated. For the noninferiority tests at week 26 and week 52, if the lower bound of the 95% CI is greater than −0.20, the avacopan group will be considered not inferior to the control group. For the superiority test, if the lower bound of the 95% CI is greater than 0.0, the avacopan group will be considered superior to the control group.

The two primary endpoints will be tested sequentially using a gatekeeping procedure to preserve the type I error rate at 0.05. The sequence of testing will be as follows:

1. Test for noninferiority of the avacopan group compared with the control group regarding remission at week 26; if the *P* value for noninferiority is *P*<.05, proceed to step 2.
2. Test for noninferiority of the avacopan group compared with the control group regarding sustained remission at week 52; if the *P* value for noninferiority is *P*<.05, proceed to step 3.
3. Test for superiority of the avacopan group compared with the control group regarding sustained remission at week 52; if the *P* value for superiority is *P*<.05, proceed to step 4.
4. Test for superiority of the avacopan group compared with the control group regarding remission at week 26.

Continuous variables of the secondary efficacy endpoints will be analyzed using a mixed effect model for repeated measures with treatment group, visit, treatment-by-visit interaction, and randomization strata as factors, and baseline as the covariate. Patients will be considered as repeated measure units over visits.

Ethics and Dissemination

Each study site is required to obtain prior approval from an Institutional Review Board on both the protocol and patient informed consent before study initiation and in accordance with International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use guidelines and the Declaration of Helsinki (amended by the 59th World Medical Association General Assembly, October 2008). Safety data and efficacy outcomes data are summarized and reviewed by an unblinded Data Monitoring Committee, regularly over the course of the study.

Patient and Public Involvement

Patients or the public were not involved in the *design* of this trial. However, patients, through a collaboration with the Vasculitis Foundation, the leading international vasculitis patient advocacy group, were involved in increasing awareness of the trial (*dissemination*) within the vasculitis patient community. It is anticipated that patients will be involved disseminating the results of this research to other patients with vasculitis, once the trial is complete and the results are made public.

Results

The last patient was enrolled in this trial in Q4 2018. Topline results are anticipated to be published by Q3 2020.

Discussion

Overview

The current therapy for ANCA-associated vasculitis improves patients' outcomes in terms of signs and symptoms of active vasculitis, but the burden of disease remains high with significant comorbidity and damage because of both glucocorticoids and disease. There is a role for targeted and less toxic treatment regimens [35]. There is increasing evidence that complement activation appears to play an important role in the pathogenesis of ANCA-associated vasculitis whereby C5aR activation augments neutrophil priming by ANCA, and directly acts as a key mediator of organ damage [4,36]. Promising results from two phase 2 studies of 12-week duration in anti-MPO- and anti-PR3-positive ANCA-associated vasculitis suggest that avacopan is well tolerated [29,30] and that its use in combination with either cyclophosphamide or rituximab may result in comparable clinical effectiveness (as assessed by the BVAS) with that of the standard-of-care glucocorticoid-containing regimen, with more rapid onset of disease control measured at week 4, as well as improved renal outcomes and quality of life [29]. Building on these observations, this larger phase 3 trial, with 1-year treatment duration and observation, aims to provide evidence for the effectiveness of C5aR inhibition for the treatment of ANCA-associated vasculitis, with potential to fundamentally change the treatment paradigm of MPA and GPA.

Study Design Features

Given the heterogeneous clinical presentation of patients with ANCA-associated vasculitis, a large sample size is planned for this study to provide an adequate number of subjects across subgroups. Moreover, the 2 treatment groups are further balanced through stratification according to three key factors that could potentially influence patient outcomes, including ANCA subtype, new or relapsing disease, and baseline immunosuppressive therapy.

This study includes a broad range of newly diagnosed or relapsing patients with ANCA-associated vasculitis. The use, if warranted, in both treatment groups, of a limited amount of nonprotocol-specified glucocorticoids is allowed at baseline, as is the use of *rescue* glucocorticoids during the study. However, the protocol specifies the indication, dose, duration, and tapering schedule for nonprotocol-supplied glucocorticoids to limit the impact of their use on the study results.

The study incorporates frequent monitoring over the 1-year treatment period, along with a pharmacodynamic sampling program. Exploratory analyses may provide information on factors that impact rates and timing to disease remission, as well as the risk of relapse among the two treatment arms.

Outcome Measures Selection

Disease Activity

This study will evaluate outcome measures at both 26 and 52 weeks along with changes over that period. The primary outcome measure of disease activity (remission) will be assessed using the BVAS version 3, the most widely accepted validated measure of disease activity and part of the Outcome Measures in Rheumatology core set of outcome measures for use in clinical trials of ANCA-associated vasculitis [37]. In this study, a team of blinded experts will adjudicate BVAS results reported by the investigators to determine if remission and sustained remission, as defined in the protocol, are attained. Slight modifications were made to the BVAS scoring criteria to suit this trial. First, all items will be scored as "new/worse" whenever they appear during the trial. Second, chronic active items will not be scored as "persistent" (persistent option removed) but are scored as per other items of activity. Third, "red blood cell casts and/or glomerulonephritis" were specified as "Other" to prompt the investigator. Fourth, for *only* the week 4 BVAS assessments, items active in the previous 7 days will be scored, not the usual 28 days.

Disease Impact

A major objective of this study will be to show the benefit of a treatment regimen for ANCA-associated vasculitis with greatly reduced cumulative dose of glucocorticoids. To assess the impact of reducing glucocorticoids, a glucocorticoid toxicity index will be used [33]. The Vasculitis Damage Index is used to document damage arising from disease and from treatment [37]. Activity of renal disease and preservation of renal function will be assessed by renal function measurements (glomerular filtration rate, urinary albumin:creatinine ratio, and urinary MCP-1:creatinine ratio).

Both clinical practice and trial data consistently demonstrate that HRQoL is impaired among patients with ANCA-associated vasculitis [37]. In a phase 2 study, HRQoL, including physical and mental health, was improved with avacopan use versus control [29]. The same instruments are used in this phase 3 study to ascertain any potential treatment differences in HRQoL. It is anticipated that the results from this study will contribute significantly to the understanding of HRQoL in relation to treatment for this patient population.

Acknowledgments

ChemoCentryx, Inc, is funding the Avacopan Development in Vasculitis to Obtain Corticosteroid elimination and Therapeutic Efficacy trial (ADVOCATE). The authors wish to thank the patients, caregivers, and the families for participating in this study. Gerri Smoluk, PhD, provided medical writing assistance in the preparation of this manuscript.

Conflicts of Interest

PM has received research grants and/or consulting fees from AbbVie, AstraZeneca, Biogen, Boeringher-Ingelheim, Bristol-Myers Squibb, Celgene, ChemoCentryx, Genentech/Roche, GlaxoSmithKline, InflaRx, Insmed, Janssen, Kiniksa, and Sanofi; DJ has received research grants and consulting fees from ChemoCentryx, Roche/Genentech, and GlaxoSmithKline; CW has been a paid consultant of ChemoCentryx; PB and JH have been employees, consultants, and shareholders of ChemoCentryx.

Multimedia Appendix 1
Avacopan Development in Vasculitis to Obtain Corticosteroid Elimination and Therapeutic Efficacy trial design summary.
[PDF File (Adobe PDF File), 135 KB - resprot_v9i4e16664_app1.pdf]

References

1. Halbwachs L, Lesavre P. Endothelium-neutrophil interactions in ANCA-associated diseases. J Am Soc Nephrol 2012 Sep;23(9):1449-1461 [FREE Full text] [doi: 10.1681/ASN.2012020119] [Medline: 22942199]
2. Furuta S, Jayne DR. Antineutrophil cytoplasm antibody-associated vasculitis: recent developments. Kidney Int 2013 Aug;84(2):244-249 [FREE Full text] [doi: 10.1038/ki.2013.24] [Medline: 23423257]
3. Kettritz R. With complements from ANCA mice. J Am Soc Nephrol 2014 Feb;25(2):207-209 [FREE Full text] [doi: 10.1681/ASN.2013101043] [Medline: 24179172]
4. Jennette JC, Falk RJ. Pathogenesis of antineutrophil cytoplasmic autoantibody-mediated disease. Nat Rev Rheumatol 2014 Aug;10(8):463-473. [doi: 10.1038/nrrheum.2014.103] [Medline: 25003769]
5. Hammerschmidt DE, Harris PD, Wayland JH, Craddock PR, Jacob HS. Complement-induced granulocyte aggregation in vivo. Am J Pathol 1981 Feb;102(2):146-150 [FREE Full text] [Medline: 7468764]
6. Tse WY, Nash GB, Hewins P, Savage CO, Adu D. ANCA-induced neutrophil F-actin polymerization: implications for microvascular inflammation. Kidney Int 2005 Jan;67(1):130-139 [FREE Full text] [doi: 10.1111/j.1523-1755.2005.00063.x] [Medline: 15610236]
7. Foreman KE, Vaporciyan AA, Bonish BK, Jones ML, Johnson KJ, Glovsky MM, et al. C5a-induced expression of P-selectin in endothelial cells. J Clin Invest 1994 Sep;94(3):1147-1155 [FREE Full text] [doi: 10.1172/JCI117430] [Medline: 7521884]

8. Schraufstatter IU, Trieu K, Sikora L, Sriramarao P, DiScipio R. Complement c3a and c5a induce different signal transduction cascades in endothelial cells. J Immunol 2002 Aug 15;169(4):2102-2110 [FREE Full text] [doi: 10.4049/jimmunol.169.4.2102] [Medline: 12165538]
9. Jones RB, Tervaert JW, Hauser T, Luqmani R, Morgan MD, Peh CA, European Vasculitis Study Group. Rituximab versus cyclophosphamide in ANCA-associated renal vasculitis. N Engl J Med 2010 Jul 15;363(3):211-220. [doi: 10.1056/NEJMoa0909169] [Medline: 20647198]
10. Stone JH, Merkel PA, Spiera R, Seo P, Langford CA, Hoffman GS, RAVE-ITN Research Group. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. N Engl J Med 2010 Jul 15;363(3):221-232 [FREE Full text] [doi: 10.1056/NEJMoa0909905] [Medline: 20647199]
11. Yates M, Watts RA, Bajema IM, Cid MC, Crestani B, Hauser T, et al. EULAR/ERA-EDTA recommendations for the management of ANCA-associated vasculitis. Ann Rheum Dis 2016 Sep;75(9):1583-1594 [FREE Full text] [doi: 10.1136/annrheumdis-2016-209133] [Medline: 27338776]
12. Little MA, Nightingale P, Verburgh CA, Hauser T, de Groot K, Savage C, European Vasculitis Study (EUVAS) Group. Early mortality in systemic vasculitis: relative contribution of adverse events and active vasculitis. Ann Rheum Dis 2010 Jun;69(6):1036-1043. [doi: 10.1136/ard.2009.109389] [Medline: 19574233]
13. Luqmani R, Suppiah R, Edwards CJ, Phillip R, Maskell J, Culliford D, et al. Mortality in Wegener's granulomatosis: a bimodal pattern. Rheumatology (Oxford) 2011 Apr;50(4):697-702. [doi: 10.1093/rheumatology/keq351] [Medline: 21112869]
14. Koldingsnes W, Nossent H. Predictors of survival and organ damage in Wegener's granulomatosis. Rheumatology (Oxford) 2002 May;41(5):572-581. [doi: 10.1093/rheumatology/41.5.572] [Medline: 12011383]
15. Corral-Gudino L, Borao-Cengotita-Bengoa M, del Pino-Montes J, Lerma-Márquez JL. Overall survival, renal survival and relapse in patients with microscopic polyangiitis: a systematic review of current evidence. Rheumatology (Oxford) 2011 Aug;50(8):1414-1423. [doi: 10.1093/rheumatology/ker112] [Medline: 21406467]
16. Takala JH, Kautiainen H, Finne P, Leirisalo-Repo M. Wegener's granulomatosis in Finland in 1981-2000: risk of dialysis-dependent renal disease. Scand J Rheumatol 2011;40(4):283-288. [doi: 10.3109/03009742.2010.533693] [Medline: 21231798]
17. Mohammad AJ, Segelmark M. A population-based study showing better renal prognosis for proteinase 3 antineutrophil cytoplasmic antibody (ANCA)-associated nephritis versus myeloperoxidase ANCA-associated nephritis. J Rheumatol 2014 Jul;41(7):1366-1373. [doi: 10.3899/jrheum.131038] [Medline: 24882836]
18. Seo P, Min Y, Holbrook JT, Hoffman GS, Merkel PA, Spiera R, WGET Research Group. Damage caused by Wegener's granulomatosis and its treatment: prospective data from the Wegener's Granulomatosis Etanercept Trial (WGET). Arthritis Rheum 2005 Jul;52(7):2168-2178 [FREE Full text] [doi: 10.1002/art.21117] [Medline: 15986348]
19. Jayne D. Evidence-based treatment of systemic vasculitis. Rheumatology (Oxford) 2000 Jun;39(6):585-595. [doi: 10.1093/rheumatology/39.6.585] [Medline: 10888702]
20. Robson JC, Dawson J, Cronholm PF, Ashdown S, Easley E, Kellom KS, et al. Patient perceptions of glucocorticoids in anti-neutrophil cytoplasmic antibody-associated vasculitis. Rheumatol Int 2018 Apr;38(4):675-682 [FREE Full text] [doi: 10.1007/s00296-017-3855-6] [Medline: 29124398]
21. Charlier C, Henegar C, Launay O, Pagnoux C, Berezné A, Bienvenu B, et al. Risk factors for major infections in Wegener granulomatosis: analysis of 113 patients. Ann Rheum Dis 2009 May;68(5):658-663. [doi: 10.1136/ard.2008.088302] [Medline: 18504289]
22. McGregor JG, Hogan SL, Hu Y, Jennette CE, Falk RJ, Nachman PH. Glucocorticoids and relapse and infection rates in anti-neutrophil cytoplasmic antibody disease. Clin J Am Soc Nephrol 2012 Feb;7(2):240-247 [FREE Full text] [doi: 10.2215/CJN.05610611] [Medline: 22134625]
23. Goupil R, Brachemi S, Nadeau-Fredette A, Déziel C, Troyanov Y, Lavergne V, et al. Lymphopenia and treatment-related infectious complications in ANCA-associated vasculitis. Clin J Am Soc Nephrol 2013 Mar;8(3):416-423 [FREE Full text] [doi: 10.2215/CJN.07300712] [Medline: 23220426]
24. Robson J, Doll H, Suppiah R, Flossmann O, Harper L, Höglund P, et al. Glucocorticoid treatment and damage in the anti-neutrophil cytoplasm antibody-associated vasculitides: long-term data from the European Vasculitis Study Group trials. Rheumatology (Oxford) 2015 Mar;54(3):471-481. [doi: 10.1093/rheumatology/keu366] [Medline: 25205825]
25. McDonough AK, Curtis JR, Saag KG. The epidemiology of glucocorticoid-associated adverse events. Curr Opin Rheumatol 2008 Mar;20(2):131-137. [doi: 10.1097/BOR.0b013e3282f51031] [Medline: 18349741]
26. Fardet L, Flahault A, Kettaneh A, Tiev KP, Généreau T, Tolédano C, et al. Corticosteroid-induced clinical adverse events: frequency, risk factors and patient's opinion. Br J Dermatol 2007 Jul;157(1):142-148. [doi: 10.1111/j.1365-2133.2007.07950.x] [Medline: 17501951]
27. Bekker P, Dairaghi D, Seitz L, Leleti M, Wang Y, Ertl L, et al. Characterization of Pharmacologic and Pharmacokinetic Properties of CCX168, a Potent and Selective Orally Administered Complement 5a Receptor Inhibitor, Based on Preclinical Evaluation and Randomized Phase 1 Clinical Study. PLoS One 2016;11(10):e0164646 [FREE Full text] [doi: 10.1371/journal.pone.0164646] [Medline: 27768695]

28. Xiao H, Dairaghi DJ, Powers JP, Ertl LS, Baumgart T, Wang Y, et al. C5a receptor (CD88) blockade protects against MPO-ANCA GN. J Am Soc Nephrol 2014 Feb;25(2):225-231 [FREE Full text] [doi: 10.1681/ASN.2013020143] [Medline: 24179165]
29. Jayne DR, Bruchfeld AN, Harper L, Schaier M, Venning MC, Hamilton P, CLEAR Study Group. Randomized Trial of C5a Receptor Inhibitor Avacopan in ANCA-Associated Vasculitis. J Am Soc Nephrol 2017 Sep;28(9):2756-2767 [FREE Full text] [doi: 10.1681/ASN.2016111179] [Medline: 28400446]
30. Merkel PA, Niles J, Jimenez R, Spiera RF, Rovin BH, Bomback A, et al. A Randomized Clinical Trial of CCX168, an Orally Administered C5aR Inhibitor for Treatment of Patients with ANCA-Associated Vasculitis. Arthritis Rheumatol 2016;68(suppl 10) [FREE Full text]
31. Mukhtyar C, Lee R, Brown D, Carruthers D, Dasgupta B, Dubey S, et al. Modification and validation of the Birmingham Vasculitis Activity Score (version 3). Ann Rheum Dis 2009 Dec;68(12):1827-1832. [doi: 10.1136/ard.2008.101279] [Medline: 19054820]
32. Suppiah R, Mukhtyar C, Flossmann O, Alberici F, Baslund B, Batra R, et al. A cross-sectional study of the Birmingham Vasculitis Activity Score version 3 in systemic vasculitis. Rheumatology (Oxford) 2011 May;50(5):899-905. [doi: 10.1093/rheumatology/keq400] [Medline: 21156667]
33. Miloslavsky EM, Naden RP, Bijlsma JW, Brogan PA, Brown ES, Brunetta P, et al. Development of a Glucocorticoid Toxicity Index (GTI) using multicriteria decision analysis. Ann Rheum Dis 2017 Mar;76(3):543-546. [doi: 10.1136/annrheumdis-2016-210002] [Medline: 27474764]
34. Exley AR, Bacon PA, Luqmani RA, Kitas GD, Gordon C, Savage CO, et al. Development and initial validation of the Vasculitis Damage Index for the standardized clinical assessment of damage in the systemic vasculitides. Arthritis Rheum 1997 Feb;40(2):371-380. [doi: 10.1002/art.1780400222] [Medline: 9041949]
35. Rodrigues JC, Walsh M. Risks and Benefits of Glucocorticoids in ANCA-Associated Vasculitis. Curr Treat Options in Rheum 2017;3(4):244-253. [doi: 10.1007/s40674-017-0081-z]
36. Kallenberg CG, Heeringa P, Stegeman CA. Mechanisms of Disease: pathogenesis and treatment of ANCA-associated vasculitides. Nat Clin Pract Rheumatol 2006 Dec;2(12):661-670. [doi: 10.1038/ncprheum0355] [Medline: 17133251]
37. Merkel PA, Aydin SZ, Boers M, Direskeneli H, Herlyn K, Seo P, et al. The OMERACT core set of outcome measures for use in clinical trials of ANCA-associated vasculitis. J Rheumatol 2011 Jul;38(7):1480-1486 [FREE Full text] [doi: 10.3899/jrheum.110276] [Medline: 21724720]

Abbreviations

ANCA: antineutrophil cytoplasmic antibody
BVAS: Birmingham Vasculitis Activity Score
EQ-5D-5L: Euro Quality of Life
GPA: granulomatosis with polyangiitis
HRQoL: health-related quality of life
MPA: microscopic polyangiitis
MPO: myeloperoxidase
PR3: proteinase 3
QD: once daily
SF-36: Short-Form 36

Edited by G Eysenbach; submitted 12.10.19; peer-reviewed by A Archer, B Johnson, G Temiz, H Demir; comments to author 01.11.19; revised version received 03.02.20; accepted 03.02.20; published 07.04.20.

Please cite as:
Merkel PA, Jayne DR, Wang C, Hillson J, Bekker P
Evaluation of the Safety and Efficacy of Avacopan, a C5a Receptor Inhibitor, in Patients With Antineutrophil Cytoplasmic Antibody–Associated Vasculitis Treated Concomitantly With Rituximab or Cyclophosphamide/Azathioprine: Protocol for a Randomized, Double-Blind, Active-Controlled, Phase 3 Trial
JMIR Res Protoc 2020;9(4):e16664
URL: https://www.researchprotocols.org/2020/4/e16664
doi:10.2196/16664
PMID:32088663

Protocol

Effects of the ACT OUT! Social Issue Theater Program on Social-Emotional Competence and Bullying in Youth and Adolescents: Protocol for a Cluster Randomized Controlled Trial

Jon Agley[1], MPH, PhD; Wasantha Jayawardene[1], MD, PhD; Mikyoung Jun[1], MPH, PhD; Daniel L Agley[1], EdD; Ruth Gassman[1], PhD; Steve Sussman[2], PhD; Yunyu Xiao[3], MPhil; Stephanie L Dickinson[4], MS

[1]Prevention Insights, Department of Applied Health Science, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States
[2]Departments of Preventive Medicine and Psychology, and School of Social Work, University of Southern California, Los Angeles, CA, United States
[3]Silver School of Social Work, New York University, New York, NY, United States
[4]Biostatistics Consulting Center, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States

Corresponding Author:
Jon Agley, MPH, PhD
Prevention Insights
Department of Applied Health Science
School of Public Health, Indiana University Bloomington
501 N Morton St
Suite 110
Bloomington, IN, 47404
United States
Phone: 1 8128551237
Email: jagley@indiana.edu

Abstract

Background: Students in the United States spend a meaningful portion of their developmental lives in school. In recent years, researchers and educators have begun to focus explicitly on social and emotional learning (SEL) in the school setting. Initial evidence from meta-analyses suggests that curricula designed to promote SEL likely produce benefits in terms of social-emotional competence (SEC) and numerous related behavioral and affective outcomes. At the same time, there are often barriers to implementing such curricula as intended, and some researchers have questioned the strength of the evaluation data from SEL programs. As part of the effort to improve programming in SEL, this paper describes the protocol for a cluster randomized trial of the ACT OUT! Social Issue Theater program, a brief psychodramatic intervention to build SEC and reduce bullying behavior in students.

Objective: The objective of this trial is to examine if a short dose of interactive psychodrama can affect SEC metrics and bullying experiences in schoolchildren in either the short (2-week) or medium (6-month) term.

Methods: The ACT OUT! trial is a cluster randomized superiority trial with 2 parallel groups. The unit of measurement is the student, and the unit of randomization is the classroom. For each grade (fourth, seventh, and 10th), an even number of classrooms will be selected from each school—half will be assigned to the intervention arm and half will be assigned to the control arm. The intervention will consist of 3 moderated psychodramatic performances by trained actors, and the control condition will be the usual school day. Outcome data will be collected at baseline (preintervention), 2-week postintervention (short term), and 6-month postintervention (medium term). Outcomes will include social-emotional competency; self-reported bullying and experiences of being bullied; receptivity to the program; and school-level data on truancy, absenteeism, and referrals to school displinary action for bullying. A power analysis adjusted for clustering effect, design effect, and potential attrition yielded a need for approximately 1594 students, consisting of an estimated 80 classrooms split evenly into intervention and control arms.

Results: This study was funded in June 2019; approved by the Indiana University Institutional review board on September 17, 2019; began subject recruitment on November 5, 2019; and prospectively registered with ClinicalTrials.gov.

Conclusions: Many states have issued recommendations for the integration of SEL into schools. The proposed study uses a rigorous methodology to determine if the ACT OUT! psychodramatic intervention is a cost-effective means of bolstering SEC and reducing bullying incidence in schools.

Trial Registration: ClinicalTrials.gov NCT04097496; https://clinicaltrials.gov/ct2/show/NCT04097496
International Registered Report Identifier (IRRID): PRR1-10.2196/17900

(JMIR Res Protoc 2020;9(4):e17900) doi:10.2196/17900

KEYWORDS

social and emotional learning; bullying; social-emotional competence; psychodrama; randomized controlled trial

Introduction

Social-Emotional Learning and Bullying

In the United States, students typically spend 1224 hours each year at school (6.8 hours per day for 180 days/year) [1], which is a substantial portion of their developmental lives. In this context, the past 25 years have seen renewed calls for the US education system to focus on holistic child development and social and emotional learning (SEL) in addition to standardized academic metrics [2]. SEL is widely considered to be important for improving students' academic and nonacademic outcomes [3]. The implementation of SEL curricula, and even the core definition of SEL, does vary somewhat [4], and researchers continue to design and test SEL programs and implementation approaches in educational settings, typically using a framework advanced by the Collaborative for Academic, Social, and Emotional Learning (CASEL), which focuses on 5 domains: self-awareness, self-management, social awareness, relationship skills, and responsible decision making [5]. Learning in these domains is expected to foster corresponding social-emotional competence (SEC) [6]. A model from a recent study (*Conceptual Model for Advancing SEL in Schools*) by Greenberg et al [7] proposed that an overall interaction between SEL and SEC domains and program implementation leads to the improvement of various short-term and long-term outcomes for students. Although individual study results have varied considerably, an examination of 4 meta-analyses of hundreds of SEL studies in the US education system found substantive evidence for short-term benefits across a wide spectrum of outcomes, including targeted SEC outcomes, improved academic performance, reduced problematic conduct frequency, and lower emotional distress. Longer-term outcomes were significant in 2 meta-analyses for academic achievement [8].

Importantly, these positive outcomes appear in some cases to extend to serious school-related behavioral issues, such as bullying. Bullying is relatively ubiquitous in US schools. The prevalence of school bullying reported in a 2014 meta-analysis of 80 studies that included data on youth aged 12-18 years was 35% for traditional bullying and 15% for cyberbullying [9]. Thus, it is notable that several studies have observed an association between SEL programs and/or constructs and the attenuation of bullying or aggressive behavior [10-12], and an inverse relationship exists between perceptions that SEL instruction is being offered and perceptions of bullying at school [13].

Although best practices for SEL education have been proposed and are being refined, there remain many barriers to implementation. These include costs related to training and teachers' time to learn about and teach SEL curricula, perceived inability to divert instructional time to SEL, potential lack of fidelity due to selective use of manualized elements, and competition with academic testing time [14,15]. For some SEL curricula, teacher time comprises a large fraction of the measured cost [16], although accurate measurement of costs and benefits of a given SEL curriculum is complex and subject to meaningful variability [17]. In addition, some scholars have been critical of the level of rigor of many SEL and SEC studies [18], asserting that the positive findings from meta-analyses may be overestimated or that the speed of SEL program adoption may exceed the generation of knowledge and understanding of how to maintain positive outcomes [19]. The current state of knowledge suggests the continued importance of research into SEL programming, both as an outcome (ie, development of SEC through SEL) and as a mediator of problem behavior (eg, bullying). To advance what is known, SEL research should prioritize methodological rigor and approaches that minimize school resource costs.

The Proposed Study

This study will be an assessment of the ACT OUT! social issue theater program as a universal SEL intervention targeting SEC and bullying in elementary, middle, and high school students. ACT OUT! is an existing program that has been performed in various forms by professionally trained members of an acting ensemble since 1995 [20]. The present iteration consists of 3 distinct 15-min scenarios per grade range (elementary, middle, and high) that present age-appropriate improvisational dramas that illustrate issues related to SEL and bullying, including facilitated discussion between the actors—who remain in character—and the students. The program lasts approximately 1 hour to fit within a typical class period within the school day, including introductions, performance, and engagement.

ACT OUT! may be contrasted with typical SEL curricula. SEL curricula tend to consist of manualized, structured classroom or multicomponent programs involving multiple sessions over time. The median number of sessions within an SEL program in a meta-analysis of 213 SEL studies was 24 [21]. Being 1 hour in duration, ACT OUT! is substantially shorter and is performed by third party, professional actors, thus meeting the goal of reduced school resource costs for SEL programming but potentially raising concerns about if such a relatively brief dose could reasonably be expected to produce an effect. Underlying this study is a supposition that unique properties of a dramatic performance may specifically trigger SEL responses. In Aristotle's *Poetics*, which is the first known work on dramatic theory, it is written that a dramatic tragedy (in the Aristotelian sense) is designed to arouse certain feelings, "wherewith to accomplish catharsis of…emotions" (as cited in an academic essay by Rosenstein [22] from a translation by Richard McKeon). This precise mechanism underlies the development

of psychodrama as a psychotherapeutic intervention, as combined action and verbalization can present a situation that elicits an emotional response "freed from the restricting stereotyped residues of past experience" [23]. Recent studies and meta-analyses have examined psychodrama as a means of prevention and/or behavior change with generally positive findings [24-30]. Researchers have also found that youth are receptive to psychodramatic elements as part of a larger prevention curriculum [31]. However, no published studies have measured any outcomes of a single-session psychodramatic SEL experience.

This will be the first study to examine if a short dose of interactive psychodrama can affect SEC metrics and bullying experiences in schoolchildren in either the short (2 weeks) or medium (6 months) term. To respond to recent criticism of SEL studies, we have chosen to utilize the SPIRIT 2013 clinical trial guidelines in developing this protocol to promote rigor, reproducibility, and transparency [32].

Methods

Ethics Approval and Consent to Participate

This study was granted approval by the Indiana University institutional review board (IRB; #1908563296).

Trial Design

The ACT OUT! trial is designed as a proof-of-concept cluster randomized superiority trial with 2 parallel groups. Although the unit of measurement is the student, the unit of randomization is the classroom, stratified by school. For each grade (fourth, seventh, and 10th), an even number of classrooms will be selected from each school; half of the selected classrooms will be randomly assigned to the intervention arm and the other half will be assigned to the control arm. If there is an odd number of classrooms for a given grade and school, the excluded classroom will be determined randomly. The use of this approach will better ensure comparable sociodemographic and school-level factors between intervention and control arms. Study hypotheses and objectives are shown in Figure 1. Participant flow, including measurement points, is shown in Figure 2.

Figure 1. Hypotheses and objectives.

1.1 Research hypothesis A
ACT OUT! Social Issue Theater is superior to TAU for development of social-emotional competence in Marion County, IN, school students enrolled in elementary, middle, and high schools.

1.2 Research hypothesis B
ACT OUT! Social Issue Theater is superior to TAU in reducing incidences of bullying in Marion County, IN, school students enrolled in elementary, middle, and high schools.

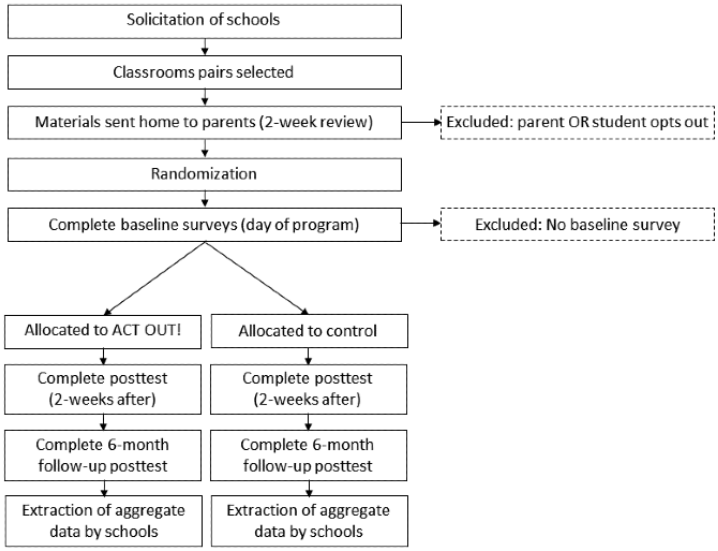
2.1 Primary objective A
To determine if ACT OUT! Social Issue Theater (1-hour exposure) is superior to TAU in developing social and emotional competence in Marion County, IN, school students enrolled in elementary, middle, and high schools, at
- 2-week posttest;
- posttest 6 months after the intervention (allowing for 2-week variability in timing).

2.2 Primary objective B
To determine if ACT OUT! Social Issue Theater (1-hour exposure) is superior to TAU in reducing bullying,
- as measured by referrals to discipline for bullying;
- self-reported bullying behavior;
- self-reported observation of bullying;

in Marion County, IN, school students enrolled in elementary, middle, and high schools, at
- 2-week posttest (selected variables);
- posttest 6 months after the intervention (allowing for 2-week variability in timing).

3.1 Secondary objectives
3.1.1 To assess student receptivity to ACT OUT! Social Issue Theater (1-hour exposure) as measured by a response scale from Dent et al (1998), and to determine whether receptivity moderates the impact of the intervention on social-emotional competence.
3.1.2 To assess targeted areas of social-emotional competence for seventh and 10th-grade students (social awareness, emotion regulation, relationship skills, and responsible decision making).
3.1.3 To conduct subanalyses of all other study outcomes by grade level (fourth, seventh, and 10th).
3.1.4 To determine if ACT OUT! Social Issue Theater (1-hour exposure) is superior to TAU in improving school-related variables,
- Truancy or absenteeism;
- Overall academic performance as measured by normalized GPA.

Figure 2. Participant timeline. GPA: grade point average; TAU: treatment as usual.



Choice of Comparators

The comparator in this study is treatment as usual (TAU), which in this case reflects continued school activity as would have otherwise occurred. This study uses a novel and low-dose intervention in a protocol that would, in a clinical setting, approximate an early phase II stage. As a result, the core questions revolve around if the intervention works at all versus if it is superior to other mechanisms that may impact the same metrics. This is consistent with recent recommendations from an expert panel from the National Institutes of Health [33].

Study Setting

The ACT OUT! trial will be conducted in public and charter schools in Marion County, Indiana, an urban county of 954,670 people who are 54.8% non-Hispanic white, 28.9% African American, and 10.6% Hispanic, with a median household income of US $44,689 [34]. The trial will recruit schools until meeting a threshold of 80 classrooms, 40 randomized to the intervention condition and 40 parallel classrooms randomized to the control condition.

Eligibility Criteria

Inclusion Criteria

Participating classrooms must comprise fourth-grade (elementary), seventh-grade (middle), or 10th-grade (high) students. Single grades within each academic level were chosen to reduce variability that might be introduced by different age and maturity ranges when interpreting effects within levels. For elementary and middle school levels, the second-highest grade level was selected. For high school, 10th grade was selected because mandatory school attendance in Indiana ends at age 16, potentially introducing bias at the level of school enrollment for grades 11 and 12.

Exclusion Criteria

If a given grade within a school has an odd number of classrooms, 1 classroom will randomly be excluded from participation. Participants and their parents or legal guardians will review the study procedures. This study will utilize a waiver of active parental consent (see Recruitment section), and so parents or guardians may opt out on behalf of their dependents, and participants may opt out themselves.

Intervention Delivery Qualifications

All individuals delivering the psychodrama interventions will be professional actors trained to the standards of Claude McNeal Productions, a professional theater company.

Intervention

Eligible classrooms will be randomized to attend a 1-hour ACT OUT! interactive, semi-improvisational psychodrama performance or to continue with their school day as normal (TAU). Interventions will be delivered within the standard school day as defined for each classroom and school. The ACT OUT! production will include 3 vignettes paired with moderated discussions between the audience and the actors, and the latter will remain in character for the duration of the intervention. Intervention guidelines will be made available alongside the published study.

Intervention Adherence

Adherence to the ACT OUT! intervention protocol will be measured at each delivery point using a fidelity checklist by a part-time employee of Claude McNeal Productions who is not a member of the acting ensemble. The fidelity checklist will contain scenario-specific items for all potential scenarios and will be made available as a multimedia appendix attached to the primary outcomes paper. Given the nature of improvisational psychodrama, the checklist will identify core areas and themes that must be addressed for the specific intervention to be considered to have been delivered with fidelity to the model. Each content area or procedure will be identified as *delivered/complete* or *undelivered/incomplete*, and the ratio will be used to generate a fidelity percentage for each delivery instance. To establish reliability, a second individual from the research team will attend 15% (6/40) of performances and complete fidelity checklists to compute interrater reliability (free-marginal kappa).

Outcomes

Primary Outcomes

Overall Social-Emotional Competence at the 2-Week and 6-Month Posttests

SEC will be measured using the scaled variable generated by the Delaware Social-Emotional Competency Scale [35]. This scale reflects 4 of CASEL's 5 core domains of SEC and has an established internal consistency coefficient of 0.84 in a large representative sample of youth. Furthermore, its developers demonstrated measurement invariance across grade levels, race, ethnicity, and gender [35]. This instrument provides an aggregate rating for overall SEC but cannot be used to generate separate scales by CASEL domains.

Self-Reported Bullying Behavior and Self-Reported Experiences of Being Bullied at the 2-Week and 6-Month Posttests

Bullying behavior and experiences of being bullied will be measured using the Bullying and Cyberbullying Scale for Adolescents (BCS-A), which uses 2 parallel 13-item scales [36]. This instrument includes items assessing physical, verbal, relational, and cyber bullying and was found in the development study to explain variance in adolescent problems beyond previously established tools—the Olweus Global Bullying Scale [37] and the Forms of Bullying Scale [38]—while displaying concurrent validity. The BCS-A also collects ratio rather than ordinal measures of bullying, which allows for face-valid use of the instrument for different time frames. As such, although the original BCS-A asked about a time frame in the past 3 months, this study will ask about a time frame in the past 2 weeks to match study protocols.

Referrals to Discipline for Bullying at the 6-Month Posttest

Referrals to discipline related to bullying will be obtained from school records in aggregate by cluster.

Secondary Outcomes

Student Receptivity to the ACT OUT! Intervention

This will be measured at the 2-week posttests using an instrument based on work published by Dent et al [39], which measures the degree to which adolescents find the event to be enjoyable, interesting, a waste of time, boring, understandable, difficult to understand, believable, important, and helpful, each on a 4-point ordinal scale.

Social-Emotional Competence Subdomains for Seventh- and 10th-Grade Students

The SEC subdomains of social awareness, emotion regulation, relationship skills, and responsible decision making will be measured using scales from the Washoe County School District Social-Emotional Competency Assessment [40]. These scales have established reliability and validity demonstrated through a unique multiyear, multimethod study, although the methods utilized did not produce traditional reliability values [40]. These items will not be collected from fourth-grade students because they would substantially increase the length of the survey instrument.

Truancy/Absenteeism and Academic Performance at 6-Months Postintervention

These will be assessed using objective data obtained from school records in aggregate by cluster.

Sample Size

The sample size was calculated based on the primary outcome (change in SEC). Notably, there are no extant studies examining the effects of professionally acted interactive psychodrama on SEC outcomes and, hence, no directly applicable effect size estimates. However, several studies have utilized psychodrama/theater as an intervention component for non-SEC outcomes, including health promotion, sexual abuse prevention, HIV prevention, violence prevention, and nutrition education. Those studies found widely varying, significant effect sizes ranging from small to large, and, in cases where effect size calculations were not included, significant improvements in outcome metrics were found. The interventions varied substantially in quality, scope, and relationship with the measured outcomes [24-29]. An earlier Cochrane review covering the years 1990 to 2006 found only 9 studies of inconsistent quality, with generally modest but positive results [30]. Separately, a robust 2012 meta-analysis by Sklad et al [41] related to SEC identified 31 universal school-based programs that addressed social learning with an aggregated Cohen *d* of 0.70 (SE of *d* 0.10), a large effect size, but those programs were longer than the ACT OUT! intervention.

On the basis of this information, we have powered this trial to detect a moderate difference in SEC (Cohen d=0.30). For a baseline sample with a parallel superiority trial, with a 2-sided significance of .05, power of 0.80, and location of the mean in 1 group as a percentile of the other group set at 62% (corresponding to Cohen d=0.30), we calculated a need for 340 participants. As this is a cluster trial, we adjusted for similarities within the groups (clustering effect) using previous research on smoking prevention [42] as a conservative metric in the absence of similar data for SEL. This yielded an intracluster correlation (ICC) value of 0.153. The design effect, assuming 20 students per classroom (m), was calculated as $[1+(m-1)\times ICC]$, yielding 3.907. The design effect multiplied by the number of participants was 1328, but we also wanted to plan for approximately 20% loss to issues related to matching and attrition. Thus, we calculated a final sample need of 1594 students, constituting an estimated 80 classrooms evenly split into intervention and control arms.

Recruitment

Each school involved in the ACT OUT! trial will be chosen based on its receptiveness and interest in participating in the project. Only schools that have provided a signed letter of agreement from an authorized official will be allowed to participate in the study. This study will utilize a waiver of parental consent (opt out), which was approved as part of the IRB submission. Neither the intervention nor the questionnaire poses more than minimal risk to participants or anyone else secondarily connected with the study (eg, their families and teachers). Surveys will be grouped by classroom and administered in a confidential manner. As part of this protocol, a notification letter and copy of the complete posttest survey

instrument will be sent home to parents a minimum of 2 weeks before initial subject interaction (parent letter included as Multimedia Appendix 1). Parents and/or guardians will be able to opt their child out of the study simply by returning the second page of the letter to the school. Students will be provided with a brief summary of the study at the top of each survey form. They will be informed that they may also opt out of the study by simply not completing the instrument or by selectively skipping items based on comfort level.

The practical purpose of using a waiver of parental consent in recruitment is the substantial difference in participation prevalence as well as potential ethical concerns resulting from the probable exclusion of underrepresented minorities and high-risk populations, as in the study by White et al [43]. Indeed, several recent studies have suggested that active parental consent and opt-in methods produce smaller student samples with different characteristics, including some related to this study (such as bullying) [43-45]. These issues mirror similar findings related to sociodemographics, risk behaviors, and study participation reported in the previous several decades of school-based research [46-49]. In general, these studies imply that active parental consent may reduce sample size, introduce bias, and thus limit the generalizability of studies with youth.

Assignment of Interventions

Allocation Sequence Generation

This trial is a cluster randomized trial; for each pair of classrooms (eg, 2 seventh-grade classrooms at the same school), a computerized random number generator will assign the classrooms to intervention (0) or control (1).

Allocation Concealment

The opt-out process will occur before the assignment. Allocation sequence will be concealed to the member of the research team who will assign each classroom to either the intervention or control group, until the moment of assignment.

Allocation Implementation

Allocation sequence generation will be completed by the research team, and data will be provided to participating schools after consent/assent has been finalized.

Blinding

Due to the nature of the ACT OUT! trial, blinding of the trial participants, school officials, and members of Claude McNeal Productions is not possible. Furthermore, as data management will occur via the primary research team, group identity cannot realistically be masked. However, we have included 2 statistical/methodological consultants in the project team who will be asked to verify all analyses using masked group assignment.

Participant Timeline

Recruitment for the study will begin in the fall of 2019, and interventions are expected to conclude in December 2019. A visual participant timeline is included in Figure 2.

Data Collection Methods

Survey Collection Procedures

Each classroom will be assigned a random cluster ID with fixed leading values for grade level and arm (eg, 4C-12345678). Each classroom will be given a study packet that contains the appropriate number of survey and response forms, a manila envelope, a white envelope, and an administrator checklist, the latter of which is available as Multimedia Appendix 2. All survey forms and manila envelopes will also be imprinted with the appropriate cluster ID.

A classroom teacher will oversee survey administration after reviewing the administrator checklist. The voluntary and confidential nature of the survey will repeatedly be emphasized per the checklist. The administrator will record the number of individuals present in the classroom on the manila envelope (serving as a denominator for the cluster response rate). Absence for the 2-week postsurvey will not prevent a participant from completing the 6-month postsurvey. All surveys distributed to students will be placed by students into the manila envelope upon completion, even if left blank, which will allow calculation of the numerator of the cluster response rate. Survey forms that were never handed out at all (*extras*) will be placed into the white envelope by the administrator.

Quality Control

Data will be collected using a form designed in Scantron Design Expert and scanned directly into a database using an Insight 700c scanner (Scantron) to avoid data entry errors associated with manual transcription.

Intervention fidelity will be recorded at each instance by an employee of Claude McNeal Productions who is not a member of the ACT OUT! ensemble. This individual will use a form generated by the researchers that documents congruence between the observed performance and a checklist of planned elements (element present/element absent). To establish reliability, a second individual from Prevention Insights will attend 15% (6/40) of performances and complete fidelity checklists to compute interrater reliability (free-marginal kappa). The checklist will be published as a multimedia appendix to the study results.

Retention

To increase individual-level participation in the study, researchers will utilize an assent procedure approved by the appropriate IRB. As noted previously, this procedure, as opposed to active parental consent, is appropriate when the risks associated with the study are minimal. Participants may, however, withdraw from the study for any reason at any time, either via parental or participant request.

As hypotheses will be tested at 2 points in time, individuals who drop out of the study after providing data at the first posttest will not be excluded from hypothesis testing for proximal (2-week) outcomes.

Data Matching

To facilitate confidentiality, researchers will need to establish longitudinal linkages between surveys without collecting

identifying information. This presents a risk of potential data loss. Meta-analytic research published within several months of trial protocol development asserts that there is no currently accepted best-practice standard for how to accomplish this most effectively in a way that preserves confidentiality and maximizes accurate matching of data between waves of collection [50]. However, these researchers found numerous benefits to accomplishing this goal using self-generated identification codes (SGICs) versus alternate methods, including true anonymity, improved response quality, utility, and maximal compliance with regulatory requirements. Researchers have also noted that SGICs are more effective when applied within smaller units or clusters [51], as is the case here.

The primary disadvantage of using an SGIC is that participants must accurately remember and report the same information [50]. The most common protocol used in longitudinal matching has been to use a combination of gender, race, and various snippets of personal information that are theoretically memorable (eg, middle initial, birth month, and mother's initial), as in the study by Kearney et al [52]. However, although such a method is generally a secure way of matching surveys without identifying individual participants, it may cause concern among parents or youth [53].

For this protocol, we opted to generate an SGIC using procedures drawn from computer science and informatics literature related to password security questions. These questions are used by both children and adults to retrieve lost passwords and are intended to be both secure and memorable [54], meeting the needs of the SGIC as identified in the survey research literature. In Rabkin's review [54] of security questions from banking institutions, questions designated as *weak* included those that are inapplicable to large segments of the target population (eg, a question about spouses), those that are not memorable (eg, last name of kindergarten teacher), and those that are ambiguous (eg, multiple truthful answers are possible, or a response is not static, such as *favorite food*) [54]. Later research on this topic also found that memorability of responses is greater for personal questions than for numbers and that memorability for all items declines over time, although high-success questions will experience minimal decline in successful recall, even after a year [55].

Using this evidence, the project will generate an SGIC with the following elements:

- Classroom ID
- Gender
- Race
- Ethnicity
- "How many *older* brothers and sisters do you have?" (0, 1, 2, 3 or more)
- "What color is your *backpack*? (if it is several colors, what color is it *mostly*?; black, red, green, blue, brown, purple, pink, a different color, I do not have a backpack)
- "What was the name of your *first pet*? (If you have never had a pet before, write the word "None" here;*handwritten entry*)
- Seventh and 10th grades only "What are the last two digits in your school locker combination? (for example, if your combination is 5-13-27, you would put 27)" (*digit entry*, I do not have a lock)

Computerized matching will occur within cluster ID. First, direct matches will be identified, and then 1-off matches (eg, all questions similar but 1) will be identified per well-established recommendations [50]. Any multiple matches (eg, more than 2 posttests matching a pretest SGIC) will be partitioned along with nonmatched data for manual inspection. Importantly, the *first pet* question includes handwritten entry, and so obvious similarities in handwriting can be used as a secondary matching tool.

Analytic Methods

General Issues

Missing data may result from nonresponse to specific items or attrition following the baseline survey (eg, dropout following pretest but before any other analyses), and the type of missingness will be analyzed. Potential types include missing completely at random (MCAR; missingness is not related to the scores of any measured variable), missing at random (MAR; missingness is related to values of other measured variables, but not to the scores of that variable itself), or missing not at random (MNAR; missingness is related to scores of that variable itself). Data that are MCAR and MAR will be managed using multiple imputation [56]. As we expect some uncertainty in matching students across time points, each baseline survey will be considered a primary key, and surveys at postintervention will only be matched with each students' baseline survey if researchers are reasonably certain that the ID code is a match from prespecified criteria. Surveys from postintervention that cannot be matched will not be included in the analysis. Students missing postintervention surveys will be considered missing due to attrition, and they will be included in the multiple imputation analysis as intention to treat.

Sensitivity analyses will be performed for plausible cases of data MNAR to examine the consistency of results across models. If the results are consistent, we will conclude that the conclusions are not compromised. We will also evaluate all appropriate statistical assumptions, such as outliers, variance heterogeneity, specification error, and nonnormality, before analysis.

Statistical Analysis

The intervention arm (ACT OUT! participation) will be compared against the control arm (TAU) in testing the primary hypotheses using an intention-to-treat model. This means that all individuals who are randomized will be included in the main analyses. As shown in Table 1, linear mixed models (LMM) or generalized LMM will be used to test all hypotheses. Baseline values will be included as covariates in each LMM, and all outcome values except receptivity and school-level data (truancy, absenteeism, and referrals to discipline for bullying) will be captured at baseline. Both fixed and random effects will be evaluated.

Table 1. Hypotheses, measures, and methods of analysis.

Outcome type and hypothesis	Time frame	Outcome measure	Method of analysis
Primary			
Intervention improves short-term SEC[a]; long-term SEC	2-week follow-up; 6-month follow-up	DSECS-S[b] (continuous)	LMM[c]; GLMM[d] if nonnormal data
Intervention reduces incidence of bullying others	2-week follow-up; 6-month follow-up	Thomas et al [36] 13-item bullying others scale (continuous)	LMM; GLMM if non-normal data
Intervention reduces incidence of being bullied	2-week follow-up; 6-month follow-up	Thomas et al [36] 13-item bullying scale (continuous)	LMM; GLMM if non-normal data
Intervention reduces incidence of referrals to discipline for bullying	6-month follow-up	Objective count data (continuous)	GLMM
Secondary			
Intervention improves social awareness (seventh and 10th grades only)	2-week follow-up; 6-month follow-up	WCSD[e] social-emotional competency assessment (continuous)	LMM; GLMM if non-normal data
Intervention improves emotion regulation (seventh and 10th grades only)	2-week follow-up; 6-month follow-up	WCSD social-emotional competency assessment (continuous)	LMM; GLMM if non-normal data
Intervention improves relationship skills (seventh and 10th grades only)	2-week follow-up; 6-month follow-up	WCSD social-emotional competency assessment (continuous)	LMM; GLMM if nonnormal data
Intervention improves responsible decision making (seventh and 10th grades only)	2-week follow-up; 6-month follow-up	WCSD social-emotional competency assessment (continuous)	LMM; GLMM if non-normal data
Receptivity to ACT OUT! will moderate impact of the intervention	2-week follow-up; 6-month follow-up	Dent et al [39] instrument (continuous)	LMM; GLMM if non-normal data
Explore outcome differences by grade level (fourth, seventh, and 10th grades)	2-week follow-up; 6-month follow-up	Four outcome measures from primary objectives (all continuous)	LMM; GLMM if non-normal data
Intervention will improve truancy and absenteeism	6-month follow-up	Objective count data (continuous)	GLMM
Intervention will improve academic performance	6-month follow-up	GPA[f] standardized to a 4.0 scale (continuous)	LMM; GLMM if non-normal data

[a]SEC: social-emotional competence.

[b]DSECS-S: Delaware Social-Emotional Competency Scale.

[c]LMM: linear mixed model.

[d]GLMM: generalized linear mixed model.

[e]WCSD: Washoe County School District.

[f]GPA: grade point average.

Results

This study was funded in June 2019; approved by the Indiana University IRB on September 17, 2019; began subject recruitment and data collection on November 5, 2019; and prospectively registered with ClinicalTrials.gov. This protocol paper, except this paragraph, was submitted as prepared in September 2019 to maintain transparency regarding changes between the proposed protocol and the finished study.

Discussion

Significance of the Study

SEL, SEC, and bullying potentially affect all schoolchildren; in recognition of this, states such as Indiana have convened bodies (eg, Indiana Commission on Improving the Status of Children in Indiana) and issued recommendations for the integration of SEL into schools [57]. At the same time, many barriers to effective program implementation remain [14-16], and the rigor of existing evaluation studies has been questioned [18]. Given the importance of these issues, this proposed study tests a novel, cost-effective intervention structure using a rigorous methodology, including this study protocol, which has been completed and registered before any subject enrollment.

Data Monitoring, Interim Analyses, and Auditing

This study will not have a data monitoring committee because the anticipated risks are minimal, and the active duration of the intervention is short within each cluster. Harms will nonetheless be evaluated in the highly unlikely event that they accrue in a manner attributable to this intervention. In addition, no interim analyses are planned other than following the 2 planned analytical timepoints indicated in Table 1. Data and analyses will be audited by multiple project consultants who are members of the project team but who are not part of Prevention Insights, the unit directly funded to support Claude McNeal Productions.

Limitations

The proposed study has several potential limitations. The protocol is designed to test outcomes from a low-dose

intervention, which, being 1 hour in length, fits within a standard class period and does not unduly burden school personnel. However, it is likely that multiple doses would provide a more robust effect (if an effect exists), and so, the pragmatic decision to test a single iteration of ACT OUT! for this study may contribute to a potential type II error. Furthermore, it is possible that an unexpectedly large percentage of students will be unable to be matched over time, especially at the 6-month posttest, due to the confidentiality requirements of the study. Use of an SGIC is a best-practice matching technique, but results from studies using this technique have disagreed about how to operationalize it and have produced a fairly wide range of matched percentages. However, collection of identifying data would require active parental consent, which itself has been shown to substantially reduce participation rates and to bias samples in ways that could meaningfully affect this study. Hence, use of the confidential SGIC to enable assent procedures is likely more robust than an identified matching protocol. Finally, it is possible that the treatment and control classes will display cross contamination, in which control subjects are informed by intervention subjects about the program. This can be identified at follow-up, and, more importantly, given that the emotional response to ACT OUT! is hypothesized to be an important mechanism of change, this threat is likely minimal because the control subjects will not have been exposed to the actual ACT OUT! experience.

Acknowledgments

The authors would like to thank David Tidd, MS, for his assistance with the protocol for confidentially linking survey instruments. This study was funded by Lilly Endowment Inc, grant no. 2019 0543.

Authors' Contributions

JA was the primary and corresponding author. JA, WJ, MJ, DA, RG, and SS conceptualized the protocol. WJ, MJ, YX, SS, and SD provided biostatistical and experimental design recommendations. All authors have read and approved the final manuscript.

Conflicts of Interest

Claude McNeal Productions and their representatives own the rights to the ACT OUT! social issue theater program. To manage this conflict, no representatives from Claude McNeal Productions were involved in preparing the protocol design, and they will not be involved in the conduct of the study, except for fidelity checklist completion (which will be verified for interrater reliability by a member of the research team). Furthermore, the lead researcher's (JA) subaward contains a clause that prevents Claude McNeal Productions from exerting oversight of the study protocol, findings, or reporting procedures.

Multimedia Appendix 1
Sample parent notification letter.
[DOCX File , 18 KB - resprot_v9i4e17900_app1.docx]

Multimedia Appendix 2
Classroom administration checklist.
[DOCX File , 24 KB - resprot_v9i4e17900_app2.docx]

References

1. Craw J. National Center on Education and the Economy. 2018. Statistic of the Month: How Much Time Do Students Spend in School? URL: http://ncee.org/2018/02/statistic-of-the-month-how-much-time-do-students-spend-in-school/ [accessed 2019-09-10]
2. Darling-Hammond L. Social and emotional learning: critical skills for building healthy schools. In: Durlak JA, Domitrovich CE, Weissberg RP, Gullotta TP, editors. Handbook of Social and Emotional Learning: Research and Practice. New York, USA: The Guilford Press; 2015:xi-xiii.
3. Zins J, Weissberg R, Wang M, Wahlberg H, editors. Building Academic Success on Social and Emotional Learning: What Does the Research Say?. New York, USA: Teachers College Press; 2004.
4. Hoffman DM. Reflecting on social emotional learning: a critical perspective on trends in the United States. Rev Educ Res 2017 Jan 10;79(2):533-556. [doi: 10.3102/0034654308325184]
5. Weissberg RP, Durlak JA, Domitrovich CE, Gullotta TP. Social and emotional learning: past, present,future. In: Durlak JA, Domitrovich CE, Weissberg RP, Gullotta TP, editors. Handbook of Social and Emotional Learning: Research and Practice. New York, USA: The Guilford Press; 2015:3-19.
6. Mahoney JL, Durlak JA, Weissberg RP. An update on social and emotional learning outcome research. Phi Delta Kappan 2018 Nov 26;100(4):18-23. [doi: 10.1177/0031721718815668]
7. Greenberg MT, Domitrovich CE, Weissberg RP, Durlak JA. Social and emotional learning as a public health approach to education. Future Child 2017;27(1):13-32. [doi: 10.1353/foc.2017.0001]

8. Domitrovich CE, Durlak JA, Staley KC, Weissberg RP. Social-emotional competence: an essential factor for promoting positive adjustment and reducing risk in school children. Child Dev 2017 Mar;88(2):408-416. [doi: 10.1111/cdev.12739] [Medline: 28213889]
9. Modecki KL, Minchin J, Harbaugh AG, Guerra NG, Runions KC. Bullying prevalence across contexts: a meta-analysis measuring cyber and traditional bullying. J Adolesc Health 2014 Nov;55(5):602-611. [doi: 10.1016/j.jadohealth.2014.06.007] [Medline: 25168105]
10. Portnow S, Downer JT, Brown J. Reductions in aggressive behavior within the context of a universal, social emotional learning program: classroom- and student-level mechanisms. J Sch Psychol 2018 Jun;68:38-52. [doi: 10.1016/j.jsp.2017.12.004] [Medline: 29861030]
11. Espelage DL, Low S, Polanin JR, Brown EC. The impact of a middle school program to reduce aggression, victimization, and sexual violence. J Adolesc Health 2013 Aug;53(2):180-186. [doi: 10.1016/j.jadohealth.2013.02.021] [Medline: 23643338]
12. Frey KS, Hirschstein MK, Snell JL, Edstrom LV, MacKenzie EP, Broderick CJ. Reducing playground bullying and supporting beliefs: an experimental trial of the steps to respect program. Dev Psychol 2005 May;41(3):479-490. [doi: 10.1037/0012-1649.41.3.479] [Medline: 15910156]
13. Nickerson AB, Fredrick SS, Allen KP, Jenkins LN. Social emotional learning (SEL) practices in schools: effects on perceptions of bullying victimization. J Sch Psychol 2019 Apr;73:74-88. [doi: 10.1016/j.jsp.2019.03.002] [Medline: 30961882]
14. Lawson GM, McKenzie ME, Becker KD, Selby L, Hoover SA. The core components of evidence-based social emotional learning programs. Prev Sci 2019 May;20(4):457-467. [doi: 10.1007/s11121-018-0953-y] [Medline: 30443846]
15. Bailey R, Stickle L, Brion-Meisels G, Jones SM. Re-imagining social-emotional learning: findings from a strategy-based approach. Phi Delta Kappan 2019 Jan 22;100(5):53-58. [doi: 10.1177/0031721719827549]
16. Belfield C, Bowden B, Klapp A, Levin H, Shand R, Zander S. The economic value of social and emotional learning. J Benefit Cost Analy 2015;6(3):508-544. [doi: 10.1017/bca.2015.55]
17. Turner AJ, Sutton M, Harrison M, Hennessey A, Humphrey N. Cost-effectiveness of a school-based social and emotional learning intervention: evidence from a cluster-randomised controlled trial of the promoting alternative thinking strategies curriculum. Appl Health Econ Health Policy 2020 Apr;18(2):271-285 [FREE Full text] [doi: 10.1007/s40258-019-00498-z] [Medline: 31347016]
18. Balfanz R, Whitehurst GJ. Should schools embrace social and emotional learning: debating the merits and costs. Educ Next 2019;19(3):68-74 GALE|A58912 [FREE Full text]
19. Wigelsworth M, Lendrum A, Oldfield J, Scott A, ten Bokkel I, Tate K, et al. The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: a meta-analysis. Camb J Educ 2016 Jul 15;46(3):347-376. [doi: 10.1080/0305764x.2016.1195791]
20. Claude MP. Claude McNeal Productions. 2019. Act Out Ensemble URL: https://www.claudemcnealproductions.com/act-out-ensemble/ [accessed 2019-09-11]
21. Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, Schellinger KB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. Child Dev 2011;82(1):405-432. [doi: 10.1111/j.1467-8624.2010.01564.x] [Medline: 21291449]
22. Rosenstein L. On Aristotle and thought in the drama. Crit Inq 1977 Apr;3(3):543-565. [doi: 10.1086/447905]
23. Davies MH. The origins and practice of psychodrama. Br J Psychiatry 1976 Sep;129:201-206. [doi: 10.1192/bjp.129.3.201] [Medline: 786420]
24. Krahé B, Knappert L. A group-randomized evaluation of a theatre-based sexual abuse prevention programme for primary school children in Germany. J Community Appl Soc Psychol 2009 Jul;19(4):321-329. [doi: 10.1002/casp.1009]
25. Lauby JL, la Pollo AB, Herbst JH, Painter TM, Batson H, Pierre A, et al. Preventing AIDS through live movement and sound: efficacy of a theater-based HIV prevention intervention delivered to high-risk male adolescents in juvenile justice settings. AIDS Educ Prev 2010 Oct;22(5):402-416. [doi: 10.1521/aeap.2010.22.5.402] [Medline: 20973661]
26. Lightfoot AF, Taboada A, Taggart T, Tran T, Burtaine A. 'I learned to be okay with talking about sex and safety': assessing the efficacy of a theatre-based HIV prevention approach for adolescents in North Carolina. Sex Educ 2015;15(4):348-363 [FREE Full text] [doi: 10.1080/14681811.2015.1025947] [Medline: 26300693]
27. Joronen K, Konu A, Rankin HS, Astedt-Kurki P. An evaluation of a drama program to enhance social relationships and anti-bullying at elementary school: a controlled study. Health Promot Int 2012 Mar;27(1):5-14. [doi: 10.1093/heapro/dar012] [Medline: 21385761]
28. Cheadle A, Cahill C, Schwartz PM, Edmiston J, Johnson S, Davis L, et al. Engaging youth in learning about healthful eating and active living: an evaluation of educational theater programs. J Nutr Educ Behav 2012;44(2):160-165. [doi: 10.1016/j.jneb.2011.06.005] [Medline: 22118997]
29. Belknap RA, Haglund K, Felzer H, Pruszynski J, Schneider J. A theater intervention to prevent teen dating violence for Mexican-American middle school students. J Adolesc Health 2013 Jul;53(1):62-67. [doi: 10.1016/j.jadohealth.2013.02.006] [Medline: 23583507]
30. Joronen K, Rankin SH, Astedt-Kurki P. School-based drama interventions in health promotion for children and adolescents: systematic review. J Adv Nurs 2008 Jul;63(2):116-131. [doi: 10.1111/j.1365-2648.2008.04634.x] [Medline: 18537845]

31. Sussman SY. Evaluating the efficacy of Project TND: evidence from seven research trials. In: Scheier LM, editor. Handbook of Adolescent Drug Use Prevention: Research, Intervention Strategies, and Practice. Washington, DC: American Psychological Association; 2015:159-176.
32. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. Br Med J 2013 Jan 8;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
33. Freedland KE, King AC, Ambrosius WT, Mayo-Wilson E, Mohr DC, Czajkowski SM, National Institutes of Health Office of Behavioral and Social Sciences Research Expert Panel on Comparator Selection in Behavioral and Social Science Clinical Trials. The selection of comparators for randomized controlled trials of health-related behavioral interventions: recommendations of an NIH expert panel. J Clin Epidemiol 2019 Jun;110:74-81. [doi: 10.1016/j.jclinepi.2019.02.011] [Medline: 30826377]
34. US Census Bureau. 2018. Quick Facts: Marion County, Indiana URL: https://www.census.gov/quickfacts/marioncountyindiana [accessed 2019-09-12]
35. Mantz LS, Bear GG, Yang C, Harris A. The delaware social-emotional competency scale (DSECS-S): evidence of validity and reliability. Child Ind Res 2016 Oct 29;11(1):137-157. [doi: 10.1007/s12187-016-9427-6]
36. Thomas HJ, Scott JG, Coates JM, Connor JP. Development and validation of the bullying and cyberbullying scale for adolescents: a multi-dimensional measurement model. Br J Educ Psychol 2019 Mar;89(1):75-94. [doi: 10.1111/bjep.12223] [Medline: 29726005]
37. Solberg ME, Olweus D. Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. Aggr Behav 2003 Apr 22;29(3):239-268. [doi: 10.1002/ab.10047]
38. Shaw T, Dooley JJ, Cross D, Zubrick SR, Waters S. The forms of bullying scale (FBS): validity and reliability estimates for a measure of bullying victimization and perpetration in adolescence. Psychol Assess 2013 Dec;25(4):1045-1057. [doi: 10.1037/a0032955] [Medline: 23730831]
39. Dent CW, Sussman S, Hennesy M, Galaif ER, Stacy AW, Moss M, et al. Implementation and process evaluation of a school-based drug abuse prevention program: project towards no drug abuse. J Drug Educ 1998;28(4):361-375. [doi: 10.2190/UFY9-WHXX-AFC1-RXB1] [Medline: 10097485]
40. Davidson LA, Crowder MK, Gordon RA, Domitrovich CE, Brown RD, Hayes BI. A continuous improvement approach to social and emotional competency measurement. J Appl Dev Psychol 2018 Mar;55:93-106. [doi: 10.1016/j.appdev.2017.03.002]
41. Sklad M, Diekstra R, Ritter MD, Ben J, Gravesteijn C. Effectiveness of school-based universal social, emotional, and behavioral programs: do they enhance students' development in the area of skill, behavior, and adjustment? Psychol Schs 2012 Oct 5;49(9):892-909. [doi: 10.1002/pits.21641]
42. Siddiqui O, Hedeker D, Flay BR, Hu FB. Intraclass correlation estimates in a school-based smoking prevention study. Outcome and mediating variables, by sex and ethnicity. Am J Epidemiol 1996 Aug 15;144(4):425-433. [doi: 10.1093/oxfordjournals.aje.a008945] [Medline: 8712201]
43. White VM, Hill DJ, Effendi Y. How does active parental consent influence the findings of drug-use surveys in schools? Eval Rev 2004 Jun;28(3):246-260. [doi: 10.1177/0193841X03259549] [Medline: 15130183]
44. Shaw T, Cross D, Thomas LT, Zubrick SR. Bias in student survey findings from active parental consent procedures. Br Educ Res J 2014 Mar 24;41(2):229-243. [doi: 10.1002/berj.3137]
45. Jelsma J, Burgess T, Henley L. Does the requirement of getting active consent from parents in school-based research result in a biased sample? An empirical study. J Empir Res Hum Res Ethics 2012 Dec;7(5):56-62. [doi: 10.1525/jer.2012.7.5.56] [Medline: 23324204]
46. Anderman C, Cheadle A, Curry S, Diehr P, Shultz L, Wagner E. Selection bias related to parental consent in school-based survey research. Eval Rev 1995;19(6):663-674. [doi: 10.1177/0193841X9501900604]
47. Dent CW, Galaif J, Sussman S, Stacy A, Burton D, Flay BR. Demographic, psychosocial and behavioral differences in samples of actively and passively consented adolescents. Addict Behav 1993;18(1):51-56. [doi: 10.1016/0306-4603(93)90008-w] [Medline: 8465677]
48. Kearney KA, Hopkins R, Mauss A, Weisheit R. Sampling bias resulting from a requirement for written parental consent. Public Opin Q 1983;47:96-102. [doi: 10.1086/268769]
49. Severson HH, Ary DV. Sampling bias due to consent procedures with adolescents. Addict Behav 1983;8(4):433-437. [doi: 10.1016/0306-4603(83)90046-1] [Medline: 6610283]
50. Audette LM, Hammond MS, Rochester NK. Methodological issues with coding participants in anonymous psychological longitudinal studies. Educ Psychol Meas 2020 Feb;80(1):163-185. [doi: 10.1177/0013164419843576] [Medline: 31933497]
51. Faden VB, Day NL, Windle M, Windle R, Grube JW, Molina BS, et al. Collecting longitudinal data through childhood, adolescence, and young adulthood: methodological challenges. Alcohol Clin Exp Res 2004 Feb;28(2):330-340 [FREE Full text] [doi: 10.1097/01.alc.0000113411.33088.fe] [Medline: 15112941]
52. Kearney KA, Hopkins RH, Mauss AL, Weisheit RA. Self-generated identification codes for anonymous collection of longitudinal questionnaire data. Public Opin Q 1984;48(1B):370-378. [doi: 10.1093/poq/48.1b.370] [Medline: 10265980]

53. Gassman RA, Dutta T, Agley J, Jayawardene W, Jun M. Social media outrage in response to a school-based substance use survey: qualitative analysis. J Med Internet Res 2019 Sep 12;21(9):e15298 [FREE Full text] [doi: 10.2196/15298] [Medline: 31516129]
54. Rabkin A. Personal Knowledge Questions for Fallback Authentication: Security Questions in the Era of Facebook. In: Proceedings of the 4th Symposium on Usable Privacy and Security. 2008 Presented at: SOUPS'08; July 23-25, 2008; Pittsburgh, Pennsylvania, USA p. 13-23 URL: https://dl.acm.org/citation.cfm?id=1408667 [doi: 10.1145/1408664.1408667]
55. Bonneau J, Bursztein E, Caron I, Jackson R, Williamson M. Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google. In: Proceedings of the 24th International Conference on World Wide Web. 2015 Presented at: WWW'15; May 18-22, 2005; Florence, Italy p. 141-150. [doi: 10.1145/2736277.2741691]
56. Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. J Am Med Assoc 2015 Nov 10;314(18):1966-1967 [FREE Full text] [doi: 10.1001/jama.2015.15281] [Medline: 26547468]
57. The Indiana Department of Education. 2019. District-Level Personnel for SEL and Mental Health URL: https://www.doe.in.gov/sites/default/files/sebw/cisc-sel-recommendation.pdf [accessed 2019-09-13]

Abbreviations

BCS-A: Bullying and Cyberbullying Scale for Adolescents
CASEL: Collaborative for Academic, Social, and Emotional Learning
ICC: intracluster correlation
IRB: institutional review board
LMM: linear mixed models
MAR: missing at random
MCAR: missing completely at random
MNAR: missing not at random
SEC: social-emotional competence
SEL: social and emotional learning
SGIC: self-generated identification code
TAU: treatment as usual

Edited by G Eysenbach; submitted 20.01.20; peer-reviewed by Z Niu; comments to author 11.02.20; revised version received 20.02.20; accepted 23.03.20; published 13.04.20.

Please cite as:
Agley J, Jayawardene W, Jun M, Agley DL, Gassman R, Sussman S, Xiao Y, Dickinson SL
Effects of the ACT OUT! Social Issue Theater Program on Social-Emotional Competence and Bullying in Youth and Adolescents: Protocol for a Cluster Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e17900
URL: http://www.researchprotocols.org/2020/4/e17900/
doi:10.2196/17900
PMID:32281541

Protocol

Comparative Study of Postural Garment Versus Exercises for Patients With Nonspecific Cervical Pain: Protocol for a Randomized Crossover Trial

Merce Avellanet[1*], MD, PhD[‡]; Anna Boada-Pladellorens[1*], MD; Jean-Claude Perrot[1], MD; Laura Loro[1], PT; Lidia Rodrigo Cansado[1], PT; David Monterde[2]; Josep Romagosa[3], MD; Elvira Gea[4], PharmD

[1]Rehabilitation Department, Hospital Nostra Senyora de Meritxell, Escaldes- Engordany, Andorra

[2]Department of Health, Catalan Health Institute, Govern de Catalunya, Barcelona, Spain

[3]Statistics Department, Govern d'Andorra, Andorra la Vella, Andorra

[4]Pharmacy Department, Hospital Nostra Senyora de Meritxell, Escaldes- Engordany, Andorra

[‡]Research Group on Health Sciences and Health Services - University of Andorra

[*]these authors contributed equally

Corresponding Author:
Merce Avellanet, MD, PhD
Rehabilitation Department
Hospital Nostra Senyora de Meritxell
Carrer dels Escalls
Escaldes- Engordany
Andorra
Phone: 376 871009
Email: merceavellanet@gmail.com

Abstract

Background: There is a high prevalence of work-related musculoskeletal disorders among health care professionals. Posture is an essential point to be addressed for health care professionals with musculoskeletal disorders. Cervical pain can result from several conditions. Treatment should include posture modification and home exercise.

Objective: This study aims to compare a new postural garment (Posture Plus Force; Medi, Bayreuth, Germany) with exercises for women with nonspecific cervical pain. The investigators focus on nurses and allied health professionals due to the importance of posture in work-related musculoskeletal disorders.

Methods: This randomized crossover clinical trial has a 3-month treatment sequence and a 3-month washout period. Participants will include nurses and allied health professionals 21 to 55 years of age with cervical pain. Participants are allocated at random to two intervention groups: a postural garment (Posture Plus Force) to be worn for 2 to 4 hours per day for 90 days (P+ group) and five physiotherapy sessions (20 minutes each) to learn stretching and strengthening exercises with instructions to continue at home on a daily basis for 90 days (Ex group). The participants in each group will crossover interventions after a 3-month washout period. The primary outcomes are postural control and pain intensity. A static posturography will be performed with a scan (SpinalMouse; Idiag AG, Fehraltorf, Switzerland). The visual analogue scale is a psychometric measuring instrument designed to document cervical pain severity in individual participants. The secondary outcomes are cervical pain-related disability, catastrophizing, the global perceived effect of treatment, and the evaluation of garment comfort. Physical activity is assessed with the International Physical Activity Questionnaire. Assessment of primary and secondary outcomes is performed at T0 (pre-intervention), T1 (immediately after garment fitting for P+ group), T30, T60, and T90. The same measurements are recorded after the washout period and during the second intervention following the same sequence. All patients are provided with a logbook for compliance recording, over the counter drug use, pain evaluation, and sick leave. Statistical analysis is conducted following intention-to-treat principles and the treatment effects calculated using linear mixed models.

Results: The study design has been approved by the Ethics Commission of Hospital N Sra de Meritxell, Andorra in March 2017. A total of 32 participants are already enrolled in the study. An extension of the study is planned in a Spanish university hospital to achieve a larger sample. Study results are expected to be published during 2020.

Conclusions: The Postural garment is expected to improve cervical pain by enhancing posture.

Trial Registration: ClinicalTrials.gov NCT03560492; https://clinicaltrials.gov/ct2/show/NCT03560492
International Registered Report Identifier (IRRID): DERR1-10.2196/14807

(JMIR Res Protoc 2020;9(4):e14807) doi:10.2196/14807

KEYWORDS

cervical pain; postural garment; cervical exercises; posture; musculoskeletal disorder

Introduction

Cervical Pain

There is a high prevalence of work-related musculoskeletal disorders among health care professionals. For instance, dental practitioners have reported a high prevalence of pain in the neck and shoulder region with rates in 2009 as high as 83% [1]. In our environment, and more concretely in our hospital, we found that during a 1-year period, 18% of the hospital staff were treated in the rehabilitation department, with more than 50% of cases caused by pain in the neck and back [2].

According to the Occupational Safety and Health Administration, musculoskeletal disorders affect the muscles, nerves, blood vessels, ligaments, and tendons [3]. Workers can be exposed to risk factors at work, such as lifting heavy items, bending, reaching overhead, pushing and pulling heavy loads, working in awkward body postures, and performing the same or similar tasks repetitively. In general, the more extreme the posture, the greater force required to complete the task.

In a recent study on dentists, a kinematic analysis underlined the posture in the area of the cervical and thoracic spine [4]. In addition, Freitag et al [5] demonstrated that sagittal trunk inclinations are an important factor in the evaluation of physical strain in nursing staff. All in all, posture is an essential point to be addressed for musculoskeletal disorders in health care professionals.

Cervical pain can result from several conditions. Disorders that cause axial neck pain include cervical strain, internal disc disruption syndrome, cervical facet-mediated pain, cervical “whiplash” syndrome, and myofascial pain. Treatment for neck pain is distinct from treatment for extremity pain or symptoms. Most patients with mild to moderate axial pain improve in 2 to 3 weeks. Treatment should include posture modification and home exercise [6-8].

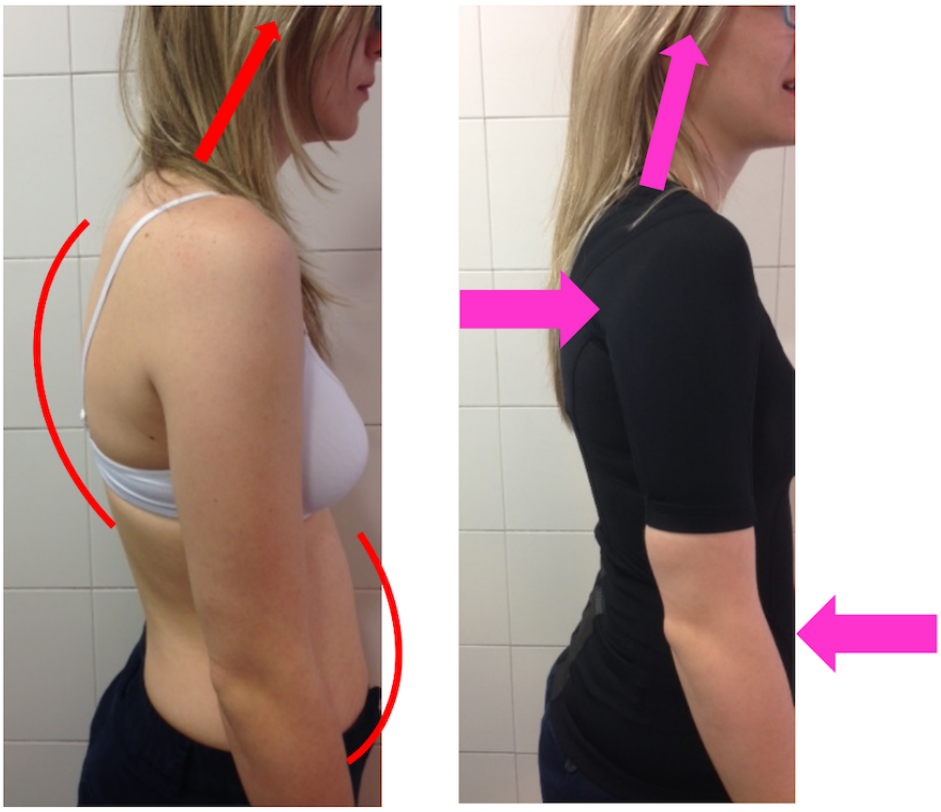
Postural Garment

The Posture Plus System adopts advanced solutions focusing on the abdomen, shoulders, and dorsal areas. The Posture Plus Force (Medi, Bayreuth, Germany) is based on the lessons learned from the development of the spinal orthosis Spinomed (Medi, Bayreuth, Germany) regarding strength and posture [9]. The Posture Plus Force is a T-shirt that includes tensional inelastic bands exercising traction for postural realignment. There are different garments available with this aim in the market. However, the Posture Plus Force is unique due to its design. Enhancement of posture is centered using an anatomical and physiological approach that focuses on abdominal and periscapular muscles.

We have tested the garment in a few patients with dorsal hyperkyphosis. We have noticed that participants reported improvements in dorsal symptoms and cervical well-being. The abdominal inelastic bands enhance core muscle contractions that facilitate a correct upright position and realign the cervical segment as well. Moreover, the dorsal bands are positioned in a functional anatomical direction following interscapular forces and enhancing perivertebral contractions instead of pulling back shoulders as other garments tend to do. Both tensional bands improve spine alignment from the abdominal zone to the cephalic position, including dorsal and cervical areas (see Figure 1). People, who do not usually remain consistent in their exercises, improved their posture and maintained a better posture during and after wearing the postural garment.

Given the high prevalence of cervical pain in health professionals previously mentioned, we hypothesized that the use of this garment can be useful for this pathology.

Figure 1. Posture Plus Force garment.



Hypothesis

The postural garment will improve cervical pain by enhancing posture.

Methods

Inclusion and exclusion criteria for our study include the following.

Inclusion Criteria

- Nurses and health allied professionals with cervical pain ≥3 on the visual analogue scale (VAS)
- Women 21 to 55 years of age
- Voluntarily signed informed consent
- Able to wear the garment and perform exercises, and attend follow-up assessments

Exclusion Criteria

- Pregnancy
- Malignancy or other severe disease
- Cervical pain with significant extremity symptoms or neurological dysfunctions (cervical radiculopathy and cervical spondylotic myelopathy)
- Unable to perform exercises
- Unwilling to do follow-up assessments
- Psychiatric disorders

Participants will be allocated at random to receive interventions in two groups based on a randomization plan.

One group will wear the postural garment (P+ group). The postural garment (Posture Plus Force) is provided to participants. It has to be worn for 2 to 4 hours per day for 90 days. See Figure 1.

The exercise group (Ex group) will do 5 physiotherapy sessions (20 minutes each) to learn stretching and strengthening exercises for the cervical and dorsal areas, with instructions to continue at home on a daily basis for 90 days. The precise explanation of exercises is shown in Multimedia Appendix 1. Home exercise maintains range of motion and helps patients become active participants in their care.

After a 3-month washout period to reverse any change that may have occurred with the first intervention, the participants will crossover the intervention for another 3-month period.

Collected Variables

The following factors will be assessed for all participants: age, weight, height, BMI, work experience, part-time employment, education, managerial position, pharmacological treatment, global perceived effect of treatment, treatment compliance, and garment comfort.

The level of physical activity is evaluated with the International Physical Activity Questionnaire (IPAQ).

Education will be divided into registered nurse, nonregistered nurse, and allied health professional. Managerial positions will include ward managers and group managers. Pharmacological treatment and sick leave will be recorded when necessary. Other variables will include main and secondary outcome measures.

Outcome Measures

The primary outcomes will be postural control and pain intensity.

Posture can be evaluated with pictures on the sagittal and frontal plane. Objective measurements of static posturography will be performed with a scan (SpinalMouse; Idiag AG, Fehraltorf, Switzerland). SpinalMouse is a device that, when combined with a PC program, assesses the curvatures of the spine without applying radiation. The device is guided manually on the skin along the spinous process apophysis from C7 to S3. The measuring head follows the contour in the sagittal plane and records clinically relevant data. A software program using a highly sophisticated algorithm uses this information to calculate the clinical parameters. The data obtained have been proven reliable and accurate [10,11] in different population groups [12-14]. The standardization and validated protocol do not include the cervical region, but dorsal kyphosis and lumbar lordosis angles influence the cervical position.

To blind the assessment of posture, the investigator that performs the SpinalMouse scan will not be aware of the participants' current intervention at each follow-up.

All other outcomes will be recorded by the other investigators.

Pain will be measured by the VAS. The VAS is a psychometric measuring instrument designed to document cervical pain severity in individual participants. It achieves a rapid (statistically measurable and reproducible) classification of symptom severity. A 100 mm horizontal line with verbal descriptors will be used to grade the amount of pain that a patient feels from no pain (left 0 mm) to an extreme amount of pain (right 100 mm).

Secondary outcomes will include cervical pain-related disability, catastrophizing, the global perceived effect of treatment, and the evaluation of garment comfort for the P+ group. Days out of work will be recorded in cases of sick leave.

Cervical pain-related disability will be assessed with the validated Spanish version of the Neck Disability Index (Multimedia Appendix 2). The scale includes 10 questions on different aspects of disability due to cervical pain [15].

Psychosocial factors are significant in predicting the duration and severity of symptoms. A significant relation between catastrophizing and pain-related outcomes has been observed in numerous pain samples [16]. Sullivan et al [17] defined catastrophizing as "an exaggerated negative mental set brought to bear during actual or anticipated painful experience". Individuals who score high on measures of pain catastrophizing report more intense pain. To assess this dimension related to pain, we will use the Pain Catastrophizing Scale (PCS) (Multimedia Appendix 3). The PCS is currently one of the most widely used measures of catastrophic thinking related to pain. It is a 13-item instrument that yields a total score and three subscale scores assessing rumination, magnification, and helplessness.

Assessment of the global perceived effect of treatment will be performed with a VAS (Multimedia Appendix 4). Garment comfort will be evaluated with a 5-point Likert scale question, an open question, and free comments (Multimedia Appendix 5).

Physical activity will be assessed with a self-reported physical activity questionnaire, the Spanish version of the IPAQ short form [18,19].

The assessment of primary and secondary outcomes will be performed at T1, T30, T60, and T90. The same measurements will be recorded after the washout period during the crossover period treatment (T0b, T30b, T60b, T90b). As previously mentioned, the same investigators will follow up with participants except for the SpinalMouse measurement, which will be performed by a blind collaborator.

All patients will be provided with a logbook for compliance recording, over the counter drug use, pain evaluation, and sick leave.

Statistical Analysis

Statistical analyses will be conducted following intention-to-treat principles, and the treatment effects will be calculated using linear mixed models. Sample size is determined with a 2 point expected difference in the VAS for pain and a sample standard deviation of 2.5.

Results

The study design has been approved by the Ethics Commission of Hospital N Sra de Meritxell, Andorra in March 2017. A total of 32 participants are already enrolled in the study. An extension of the study is planned in a Spanish university hospital to obtain a larger sample. Study results are expected to be published during 2020.

Discussion

Cervical Pain and Posture

The main component of cervical pain management is exercise. A correct posture is maintained with core muscle contractions. The Posture Plus Force postural garment may enhance posture by muscle activation. Both exercise and postural improvement are expected to improve cervical pain.

Potential Risks

No significant risks can emerge from this protocol study.

All patients will be assessed before being included and during the follow-up by a physical medicine and rehabilitation physician. Exercises will be taught by a certified physiotherapist.

Regarding the suitability of interventions, it has been stated that usual physiotherapy may be only marginally better than brief physiotherapy intervention (1-3 sessions) for neck pain [20]. According to our clinical practice, any patient will be free to schedule an earlier appointment if warranted during the study period.

Ethical Issues

The study design has been approved by the Ethics Commission of Hospital N Sra de Meritxell, Andorra. The standards of the CONSORT (Consolidated Standards of Reporting Trials) statement of 1996, as revised in 2010, will be followed.

Prior to being enrolled in the study, all patients will be informed about the study objectives, measurements and pictures to be taken. All participants have to give their consent to being photographed in advance. All the participants have to sign the informed consent before enrollment.

Financial Disclosure

No funding is received by investigators or participants.

Medi (Bayreuth, Germany) provides the logbook and postural garments for all participants.

This study has the approval of the hospital medical director. Costs will be recorded and calculated for both interventions. In our environment, typical protocol treatment for cervical pain includes 2 to 3 consultations with a physical medicine and rehabilitation physician, and 15 to 20 physiotherapy sessions. For this study protocol, assessment visits will be higher than the standard procedure, but physiotherapy sessions will dramatically diminish. Moreover, participants will aim to be an active part of their neck pain care. In conclusion, we expect a decrease in costs compared to usual cervical pain approaches in our department combined with the educational self-management of pain.

Acknowledgments

The authors would like to thank Prof Dr Michael Pfeiffer for supporting the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Exercise protocol.
[PPTX File , 782 KB - resprot_v9i4e14807_app1.pptx]

Multimedia Appendix 2
Spanish version of Neck Disability Index.
[DOCX File , 16 KB - resprot_v9i4e14807_app2.docx]

Multimedia Appendix 3
Pain Catastrophizing Scale Spanish version.
[DOCX File , 14 KB - resprot_v9i4e14807_app3.docx]

Multimedia Appendix 4
Assessment of global perceived effect of treatment.
[DOCX File , 13 KB - resprot_v9i4e14807_app4.docx]

Multimedia Appendix 5
Garment comfort evaluation.
[DOCX File , 13 KB - resprot_v9i4e14807_app5.docx]

References

1. Morse T, Bruneau H, Dussetschleger J. Musculoskeletal disorders of the neck and shoulder in the dental professions. Work 2010;35(4):419-429. [doi: 10.3233/WOR-2010-0979] [Medline: 20448321]
2. Avellanet Viladomat M, Sáenz Guerrero A, Hijós Bitrián E, Romagosa Massana J. Asistencia al personal hospitalario en un servicio de rehabilitación. ¿Cómo es y cuánto cuesta? Rehabilitación 2005 Jan;39(2):66-69. [doi: 10.1016/S0048-7120(05)74315-2]
3. Occupational Safety and Health Administration. OSHA.gov. Ergonomics URL: https://www.osha.gov/SLTC/ergonomics/index.html [accessed 2018-05-12]
4. Ohlendorf D, Erbe C, Hauck I, Nowak J, Hermanns I, Ditchen D, et al. Kinematic analysis of work-related musculoskeletal loading of trunk among dentists in Germany. BMC Musculoskelet Disord 2016 Oct 18;17(1):427 [FREE Full text] [doi: 10.1186/s12891-016-1288-0] [Medline: 27756271]
5. Freitag S, Fincke-Junod I, Seddouki R, Dulon M, Hermanns I, Kersten JF, et al. Frequent bending--an underestimated burden in nursing professions. Ann Occup Hyg 2012 Jul;56(6):697-707. [doi: 10.1093/annhyg/mes002] [Medline: 22356807]
6. Isaac Z, Atlas SJ, Dashe J. Treatment of neck pain. UpToDate 2016 Jun.

7. Gross A, Kay TM, Paquin J, Blanchette S, Lalonde P, Christie T, Cervical Overview Group. Exercises for mechanical neck disorders. Cochrane Database Syst Rev 2015 Jan 28;1:CD004250. [doi: 10.1002/14651858.CD004250.pub5] [Medline: 25629215]
8. Im B, Kim Y, Chung Y, Hwang S. Effects of scapular stabilization exercise on neck posture and muscle activation in individuals with neck pain and forward head posture. J Phys Ther Sci 2016 Mar;28(3):951-955 [FREE Full text] [doi: 10.1589/jpts.28.951] [Medline: 27134391]
9. Pfeifer M, Kohlwey L, Begerow B, Minne HW. Effects of two newly developed spinal orthoses on trunk muscle strength, posture, and quality-of-life in women with postmenopausal osteoporosis: a randomized trial. Am J Phys Med Rehabil 2011 Oct;90(10):805-815. [doi: 10.1097/PHM.0b013e31821f6df3] [Medline: 21681065]
10. Post RB, Leferink VJM. Spinal mobility: sagittal range of motion measured with the SpinalMouse, a new non-invasive device. Arch Orthop Trauma Surg 2004 Apr;124(3):187-192. [doi: 10.1007/s00402-004-0641-1] [Medline: 14968367]
11. Mannion AF, Knecht K, Balaban G, Dvorak J, Grob D. A new skin-surface device for measuring the curvature and global and segmental ranges of motion of the spine: reliability of measurements and comparison with data reviewed from the literature. Eur Spine J 2004 Mar;13(2):122-136 [FREE Full text] [doi: 10.1007/s00586-003-0618-8] [Medline: 14661104]
12. Lang-Tapia M, España-Romero V, Anelo J, Castillo MJ. Differences on spinal curvature in standing position by gender, age and weight status using a noninvasive method. J Appl Biomech 2011 May;27(2):143-150. [doi: 10.1123/jab.27.2.143] [Medline: 21576723]
13. Mizukami S, Abe Y, Tsujimoto R, Arima K, Kanagae M, Chiba G, et al. Accuracy of spinal curvature assessed by a computer-assisted device and anthropometric indicators in discriminating vertebral fractures among individuals with back pain. Osteoporos Int 2014 Jun;25(6):1727-1734. [doi: 10.1007/s00198-014-2680-y] [Medline: 24627138]
14. Livanelioglu A, Kaya F, Nabiyev V, Demirkiran G, Fırat T. The validity and reliability of "Spinal Mouse" assessment of spinal curvatures in the frontal plane in pediatric adolescent idiopathic thoraco-lumbar curves. Eur Spine J 2016 Feb;25(2):476-482. [doi: 10.1007/s00586-015-3945-7] [Medline: 25900295]
15. Ortega JA, Martínez AD, Ruiz R. Validation of the Spanish version of the Neck Disability Index. Spine (Phila Pa 1976) 2010 Feb 15;35(4):E114-E118. [doi: 10.1097/BRS.0b013e3181afea5d] [Medline: 20110848]
16. Nordin CA, Michaelson P, Gard G, Eriksson MK. Effects of the web behavior change program for activity and multimodal pain rehabilitation: randomized controlled trial. J Med Internet Res 2016 Oct 05;18(10):e265 [FREE Full text] [doi: 10.2196/jmir.5634] [Medline: 27707686]
17. Sullivan MJ, Thorn B, Haythornthwaite JA, Keefe F, Martin M, Bradley LA, et al. Theoretical perspectives on the relation between catastrophizing and pain. Clin J Pain 2001 Mar;17(1):52-64. [doi: 10.1097/00002508-200103000-00008] [Medline: 11289089]
18. Roman-Viñas B, Serra-Majem L, Hagströmer M, Ribas-Barba L, Sjöström M, Segura-Cardona R. International Physical Activity Questionnaire: reliability and validity in a Spanish population. European Journal of Sport Science 2010 Sep;10(5):297-304. [doi: 10.1080/17461390903426667]
19. Silsbury Z, Goldsmith R, Rushton A. Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. BMJ Open 2015 Sep 15;5(9):e008430 [FREE Full text] [doi: 10.1136/bmjopen-2015-008430] [Medline: 26373402]
20. Klaber Moffett JA, Jackson DA, Richmond S, Hahn S, Coulton S, Farrin A, et al. Randomised trial of a brief physiotherapy intervention compared with usual physiotherapy for neck pain patients: outcomes and patients' preference. BMJ 2005 Jan 08;330(7482):75 [FREE Full text] [doi: 10.1136/bmj.38286.493206.82] [Medline: 15585539]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
IPAQ: International Physical Activity Questionnaire
PCS: Pain Catastrophizing Scale
VAS: visual analogue scale.

Edited by G Eysenbach; submitted 29.05.19; peer-reviewed by M Alshehri, E Sadeghi-Demneh; comments to author 05.12.19; revised version received 08.12.19; accepted 14.01.20; published 16.04.20.

Please cite as:
Avellanet M, Boada-Pladellorens A, Perrot JC, Loro L, Rodrigo Cansado L, Monterde D, Romagosa J, Gea E
Comparative Study of Postural Garment Versus Exercises for Patients With Nonspecific Cervical Pain: Protocol for a Randomized Crossover Trial
JMIR Res Protoc 2020;9(4):e14807
URL: https://www.researchprotocols.org/2020/4/e14807
doi:10.2196/14807
PMID:32297876

Protocol

Hemostatis Analyzer-Supported Hemotherapy Algorithm in Cardiac Surgery: Protocol for a Randomized Controlled Monocentric Trial

Sophie Michel[1*], MD; Florian Piekarski[1*], MD; Jan-Hendrik Fischer[1], MD; Vanessa Hettler[1], MD; Elisabeth Hannah Adam[1], MD; Lars Holzer[1], MD; Gösta Lotz[1], DESA, MA; Thomas Walther[1], MD; Kai Zacharowski[1], MD, PhD, ML, FRCA; Florian Jürgen Raimann[1], MD

Department of Anaesthesiology, Intensive Care Medicine and Pain Therapy, University Hospital Frankfurt, Goethe University, Frankfurt am Main, Germany

[*]these authors contributed equally

Corresponding Author:
Florian Jürgen Raimann, MD
Department of Anaesthesiology, Intensive Care Medicine and Pain Therapy
University Hospital Frankfurt
Goethe University
Theodor-Stern-Kai 7
Haus 13
Frankfurt am Main, 60590
Germany
Phone: 49 15117190152
Email: Florian.Raimann@kgu.de

Abstract

Background: Point of care devices for performing targeted coagulation substitution in patients who are bleeding have become increasingly important in recent years. New on the market is the Quantra. It is a device that uses sonorheometry, a sonic estimation of elasticity via resonance, which is a novel ultrasound-based technology that measures viscoelastic properties of whole blood. Several studies have already shown the comparability of the Quantra with devices already established on the market, such as the rotational thromboelastometry (ROTEM) device.

Objective: In contrast to existing studies, this study is the first prospective interventional study using this new system in a cardiac surgical patient cohort. We will investigate the noninferiority between an already existing coagulation algorithm based on the ROTEM/Multiplate system and a new algorithm based on the Quantra system for the treatment of coagulopathic cardiac surgical patients.

Methods: The study is divided into two phases. In an initial observation phase, whole blood samples of 20 patients obtained at three defined time points (prior to surgery, after completion of cardiopulmonary bypass, and on arrival in the intensive care unit) will be analyzed using both the ROTEM/Multiplate and Quantra systems. The obtained threshold values will be used to develop a novel algorithm for hemotherapy. In a second intervention phase, the new algorithm will be tested for noninferiority against an algorithm used routinely for years in our department.

Results: The main objective of the examination is the cumulative loss of blood within 24 hours after surgery. Statistical calculations based on the literature and in-house data suggest that the new algorithm is not inferior if the difference in cumulative blood loss is <150 mL/24 hours.

Conclusions: Because of the comparability of the Quantra sonorheometry system with the ROTEM measurement methods, the existing hemotherapy treatment algorithm can be adapted to the Quantra device with proof of noninferiority.

Trial Registration: ClinicalTrials.gov NCT03902275; https://clinicaltrials.gov/ct2/show/NCT03902275

International Registered Report Identifier (IRRID): DERR1-10.2196/17206

(JMIR Res Protoc 2020;9(4):e17206) doi:10.2196/17206

KEYWORDS

Quantra; cardiothoracic surgery; bypass; coagulopathy; point of care; algorithm; rotational thromboelastometry; Multiplate

Introduction

Background

A targeted coagulation therapy during intra- and postoperative care for cardiac surgical patients needs an accurate knowledge of their hemostatic conditions. For the purpose of coagulation diagnosis, many institutions take blood samples and send them to the central clinical chemist. After analysis and validation, results are transmitted electronically, which might prolong the time required to derive therapeutic interventions.

In recent years, point-of-care testing (POCT) devices for the diagnosis of patients who are coagulopathic have become increasingly important. In a neurosurgical trial, Beynon et al [1] showed that the use of POCT markedly reduced the time to receive clotting parameter results in comparison to conventional laboratory analyses. Moreover, the quality of the results (eg, the international normalized ratio and conventional laboratory results) in POCT also showed a high correlation in values [2].

POCT devices have been routinely used for guiding intra- and postoperative targeted coagulation therapy for years in our clinic. The devices in use include the rotational thromboelastometry (ROTEM) delta and Multiplate. For the purpose of analysis, a whole blood sample is pipetted and mixed with test reagents. Depending on the selected reagents, different steps of the coagulation cascade can be evaluated, and, according to each parameter result, an appropriate therapy can be derived.

A coagulation algorithm based on those measurements has been developed in our clinic, and a modified version for cardiac surgery has been used successfully for years in perioperative coagulation management [3]. The algorithm requires additional information on the platelet function, which is also carried out as a standard practice on the bedside using the Multiplate device.

The Quantra Analyzer system from HemoSonics, a new system for hemostasis analysis, has recently become available on the market [4]. The Quantra system also allows the analysis of a whole blood sample on the bedside. Moreover, due to the fully sealed cartridge system of the Quantra, pipetting of a whole blood sample is no longer necessary and hence avoids the time-consuming and potentially error-prone procedure. In addition, the risk of infection for the examiner is smaller. An evaluation of the platelet activity for the Quantra device system in a cardiac surgical patient trial, [5] as well as the comparability procedures with ROTEM measures has already been assessed [5,6].

Objectives

In this first interventional study, our main objective is to show the noninferiority of a new Quantra-based hemotherapy algorithm in comparison to an existing algorithm based on the ROTEM delta and Multiplate used in our clinical routine. We expect comparable results in the effects of stabilizing coagulation after cardiosurgical interventions and postoperative blood loss.

Methods

Materials

To guarantee effective coagulation management during cardiosurgical interventions, we already use a coagulation algorithm adapted to the results of point-of-care measurements for coagulation and platelet function.

In our clinic the ROTEM device is used to analyze clotting time, clot formation, clot stiffness, and the dissolution of the clot (fibrinolysis). In addition to ROTEM, we use the Multiplate system to obtain information about the platelet function and to detect a possible presence of adenosine-diphosphate (ADP) antagonists like Clopidogrel (ADPtest) or cyclooxygenase-inhibitors (arachiconic acid test). Based on the results of the ROTEM and Multiplate system, we administer a systematic coagulation therapy in accordance with an already existing algorithm.

A new system, the Quantra Hemostasis Analyzer (Figure 1) is based on sonic estimation of elasticity via resonance. Sonorheometric technology uses high-frequency ultrasound for analyzing changes in viscoelastic properties of whole blood samples (Figure 2). For this purpose, a cartridge with four channels, containing different reagents in every channel, is used (see Table 1).

Figure 1. Quantra Analyzer. CS: clot stiffness; CT: clot time; CTH: heparinase clot time; CTR: clot time ratio; FCS: fibrinogen contributions to clot stiffness; PCS: platelet contributions to clot stiffness.

Figure 2. Principle of SEER Sonorheometry.

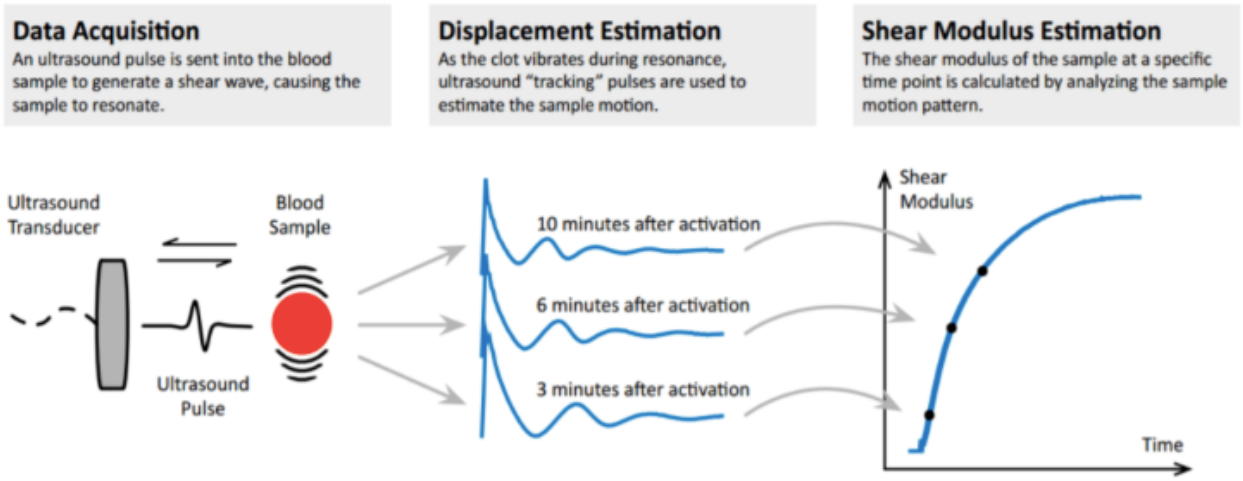


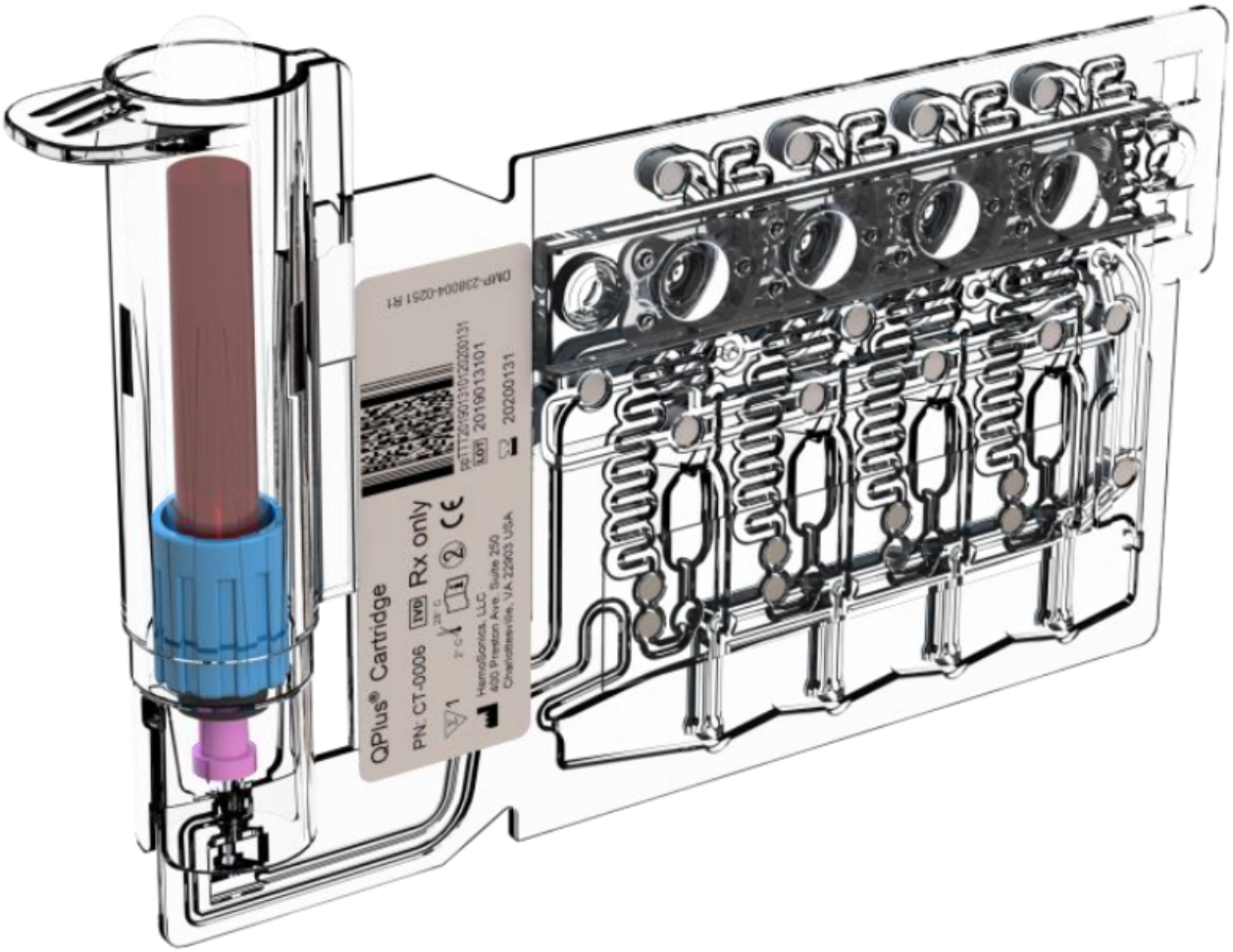
Table 1. Design and assignment of the individual channels of the measuring cartridge.

Cartridge channel	Reagents	Possible measurements
1	• Kaolin (intrinsic coagulation)	Clot time
2	• Kaolin • Heparinase (heparin neutralization)	Clot time with heparinase
3	• Thromboplastin (extrinsic coagulation, tissue factor activation) • Polybrene (heparin neutralization)	Clot stiffness
4	• Thromboplastin (tissue factor activation) • Polybrene (heparin neutralization) • Abciximab (platelet inhibition)	Fibrinogen contribution
1-2	N/A[a]	Clot time ratio
3-4	N/A	Platelet contribution

[a]N/A: not applicable.

The user inserts a whole blood sample (collected in a 2.7 ml sodium citrate tube) into the cartridge (QPlus cartridge, Figure 3). The four tests run simultaneously and in a self-contained system. An ultrasound pulse, which is transmitted into the blood sample chamber, generates a shear wave as it clots. Thus, the resonance generated by the displacement of a developing clot in the blood sample is recorded. A series of ultrasound pulses is necessary to estimate the sample's motion. The generated shear wave modulus (stiffness) of the blood is then measured at precise time points [7].

Figure 3. QPlus cartridge.

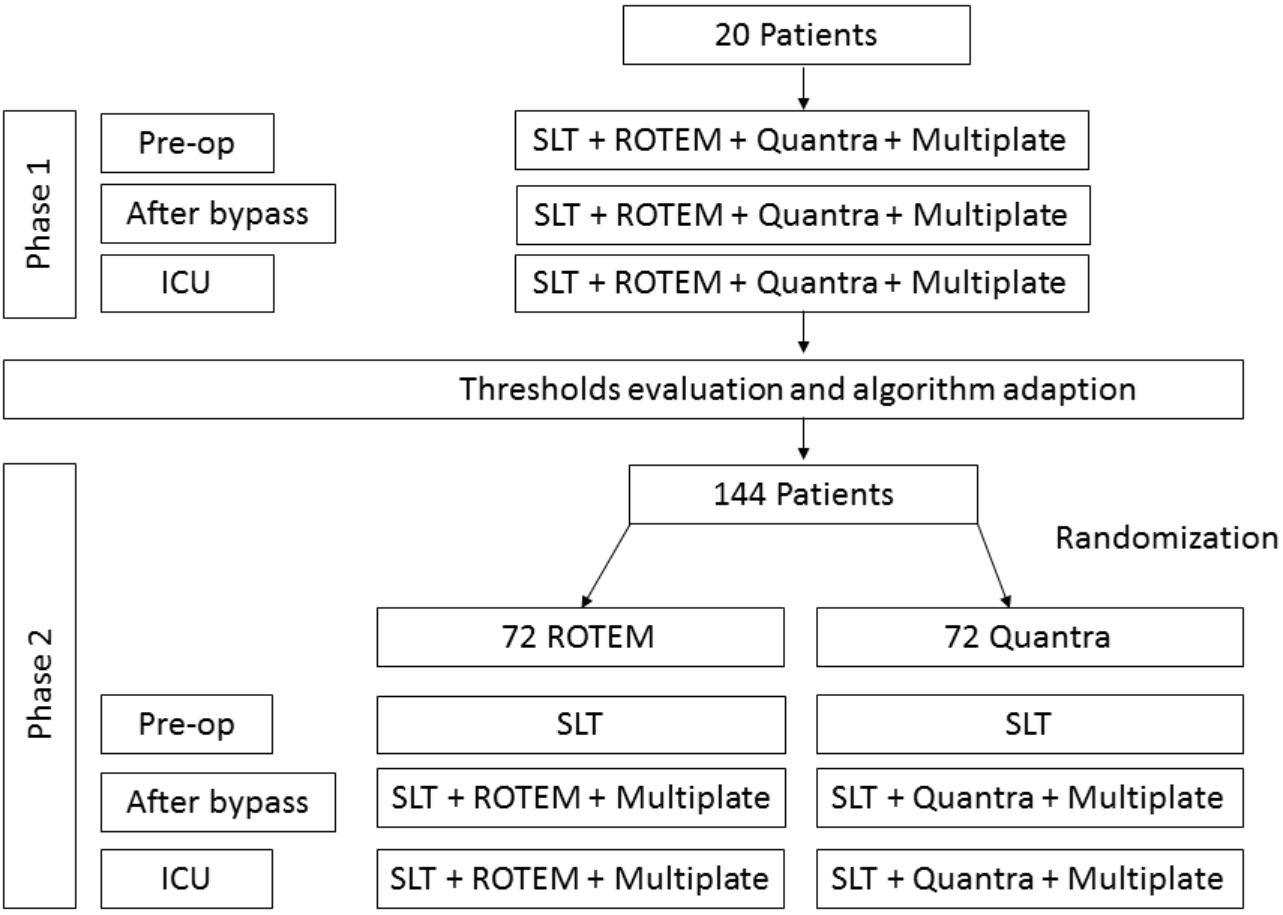


The Quantra system has already been used for evaluation of hemostatic function in critical care and operative settings, and was compared with laboratory parameters and ROTEM [5,7].

Study Phases

The study is divided into two phases (see Figure 4). Recruiting started in September 2019. It is planned to include 20 patients in phase one and 144 patients in phase two. The estimated end of the study is calculated to be December 2020.

Figure 4. Study flow chart. ICU: intensive care unit; ROTEM: rotational thromboelastometry; SLT: standard laboratory test.



Patient Population

Inclusion criteria will be patients 18 years of age or older undergoing elective cardiac surgery on cardiopulmonary bypass.

Patients without written consent or missing data will be excluded.

Observational Phase

In an initial phase, 20 patients undergoing elective cardiac surgery will be examined at the following three time points, using the ROTEM and Multiplate, as well as the Quantra.

1. Directly prior to surgery
2. After completion of the bypass and before reversal of heparin with protamine
3. On arrival in the intensive care unit

The data obtained will be used to evaluate results and threshold values for therapy. Those results will serve to review and adapt the already existing algorithm used in our clinic. All therapeutic decisions at this time will still be made based on the existing coagulation algorithm using the ROTEM and Multiplate results.

Interventional Phase

After adapting the existing algorithm to the Quantra POCT results, the second phase will follow to evaluate the use of the new algorithm. For this purpose, 144 patients will be randomized into two groups. Patients of the Quantra group (QG) will be treated according to the new algorithm, while the comparison, or standard group (SG), will be treated using the existing algorithm. All relevant data will be entered in a case report form. Table 2 gives an overview of the tests within the various phases of the study compared with the measurements commonly used in our clinic before, during, and after surgery.

Table 2. Planned tests during the study compared to the previous standard in our clinic.

Phase	Standard care	Study procedures	
		Quantra group	Standard group
Phase 1			
Presurgery	• Standard laboratory test[a]	• Standard laboratory test • Quantra/ROTEM[b]/Multiplate	• Standard laboratory test • Quantra/ROTEM/Multiplate
During surgery	• ROTEM/Multiplate only when bypass >120 min	• Standard laboratory test • Quantra/ROTEM/Multiplate	• Standard laboratory test • Quantra/ROTEM/Multiplate
Postsurgery	• Standard laboratory test • ROTEM/Multiplate only if persistent bleeding	• Standard laboratory test • Quantra/ROTEM/Multiplate	• Standard laboratory test • Quantra/ROTEM/Multiplate
Phase 2			
Presurgery	• Standard laboratory test	• Standard laboratory test	• Standard laboratory test
During surgery	• ROTEM/Multiplate only when bypass >120 min	• Standard laboratory test • Quantra/Multiplate	• Standard laboratory test • ROTEM/Multiplate
Postsurgery	• Standard laboratory test • ROTEM/Multiplate only in case of persistent bleeding	• Standard laboratory test • Quantra/Multiplate	• Standard laboratory test • ROTEM/Multiplate

[a]Standard laboratory tests include hemoglobin, thrombocytes, international normalized ratio, Quick test, activated partial thromboplastin time, antithrombine III, fibrinogen, creatinine, and electrolytes.

[b]ROTEM: rotational thromboelastometry.

Statistical Analysis

As this is a pilot study, sample size was calculated based on the existing literature and our own clinical data. Our own data suggest that moderate blood loss of 650 mL is expected within 24 hours after surgery.

Sample Size Consideration

To determine the required number of cases for a statistical power of 80%, it was assumed that the blood loss in both groups would be normally distributed. In-house data shows that blood loss is, for obvious reasons, not normally distributed, and a difference of 150 mL corresponds to a Mann-Whitney estimator of 0.363. Therefore, the one-sided noninferiority test requires a total of 140 evaluable patients (70 in each group) to obtain a statistical power of 80%. To compensate for 2 patients per group being lost to follow-up each groups' target size is 72 patients. This results in a total population of 144 patients.

Primary Outcome Analysis

A one-sided significance level of alpha 2.5% will be used in proving the noninferiority of the QG compared to the SG. The noninferiority limit has been set to 150 mL. For testing, a nonparametric Wilcoxon-Mann-Whitney U test will be used for one-sided tests of superiority. In the initial Quantra group 150 mL will be subtracted from the primary target test, as this corresponds to a nonparametric variant of a usual noninferiority test with the Wilcoxon-Mann-Whitney U test.

Conversions

Quantra clot stiffness values are expressed in hectopascals, whereas corresponding ROTEM values are expressed as an amplitude (A) in mm. The relationship between A in mm and shear modulus (G) in pascals is not linear. For a proper comparison, the ROTEM A (mm) needs to be converted to pascals by using the following formula: G (pascals) = (500 x A) / (100 – A), as described by Solomon C et al [8].

Results

The main and secondary outcomes are shown in Textbox 1.

The primary outcome will be the cumulative blood loss after cardiac surgery (coronary bypass, valve, and aortic operations) within 24 hours.

Textbox 1. Main and Secondary Outcomes.

Primary outcome

- Cumulative blood loss after cardiac surgery (24 hours)

Secondary outcomes

- Correlation between laboratory and point-of-care testing results (Quick test, international normalized ratio, activated partial thromboplastin time, fibrinogen)
- Platelet function with Multiplate-Analysis
- Time until results are obtained (ie, time until therapy can be established)
- Use of blood products (packed red blood cells, fresh frozen plasma, platelets)
- Thromboembolic events
- Comorbidities
- Type of surgery procedure
- Demographic data
- Use of procoagulants (fibrinogen, desmopressin, tranexamic acid, prothrombin complex concentrate, recombinant enabled factor VIIa, calcium)

We expect to receive comparable results regarding the named main and secondary outcomes, both in the SG as well as in the QG. According to this assumption, we want to show the noninferiority of a Quantra-based hemotherapy algorithm compared to the ROTEM and Multiplate system.

This study is approved by the local ethical committee (ID: 42/19) and was started in September 2019. Results will be published soon after completion.

Discussion

This is the first prospective interventional study comparing the Quantra system with the ROTEM and Multiplate system regarding blood loss in 24 hours after implementing an adapted coagulation algorithm for cardiac surgical patients. Huffmeyer et al [5] showed a significant correlation between the Quantra system and the ROTEM and Multiplate system with respect to clot stiffness and fibrinogen contribution [5].

The focus of our study is whether a Quantra-based hemotherapy algorithm is noninferior to the ROTEM and Multiplate-based algorithm we already use in our clinical routine.

There are some aspects of the Quantra system that could prove beneficial. First, there is no need for pipetting reagents to the full blood sample. Accordingly, we expect less erroneous measurements due to inaccurate pipetting.

Both the accuracy of the results and the time until results are available correlate with the skills of the investigator. We expect the Quantra system to be easy to use and interpret. There will be no extensive training required.

Second, the Quantra Analyzer is a closed system, which minimizes the probability of contamination. In comparison, the ROTEM and Multiplate is an open system and possibly more sensitive to contamination.

Acknowledgments

The study has been carried out in cooperation with HemoSonics. The principal investigator received two devices for testing and the necessary disposables. However, the company had no influence on the study protocol or decision to publish.

Conflicts of Interest

Author FR received speaking honoraria from Keller Medical.

References

1. Beynon C, Jakobs M, Rizos T, Unterberg AW, Sakowitz OW. Rapid bedside coagulometry prior to urgent neurosurgical procedures in anticoagulated patients. Br J Neurosurg 2014 Jan 09;28(1):29-33. [doi: 10.3109/02688697.2013.869549] [Medline: 24313307]
2. Beynon C, Erk AG, Potzy A, Mohr S, Popp E. Point of care coagulometry in prehospital emergency care: an observational study. Scand J Trauma Resusc Emerg Med 2015 Aug 12;23(1):58 [FREE Full text] [doi: 10.1186/s13049-015-0139-6] [Medline: 26260487]
3. Weber CF, Görlinger K, Meininger D, Herrmann E, Bingold T, Moritz A, et al. Point-of-care testing: a prospective, randomized clinical trial of efficacy in coagulopathic cardiac surgery patients. Anesthesiology 2012;117(3):531-547. [doi: 10.1097/aln.0b013e318264c644]

4. Viola F, Lin-Schmidt X, Bhamidipati C, Haverstick DM, Walker WF, Ailawadi G, et al. Sonorheometry assessment of platelet function in cardiopulmonary bypass patients: correlation of blood clot stiffness with platelet integrin αIIbβ3 activity, aspirin usage, and transfusion risk. Thromb Res 2016 Feb;138:96-102. [doi: 10.1016/j.thromres.2015.11.036] [Medline: 26688324]
5. Huffmyer JL, Fernandez LG, Haghighian C, Terkawi AS, Groves DS. Comparison of SEER sonorheometry with rotational thromboelastometry and laboratory parameters in cardiac surgery. Anesthesia & Analgesia 2016;123(6):1390-1399. [doi: 10.1213/ane.0000000000001507]
6. Reynolds PS, Middleton P, McCarthy H, Spiess BD. A comparison of a new ultrasound-based whole blood viscoelastic test (SEER sonorheometry) versus thromboelastography in cardiac surgery. Anesthesia & Analgesia 2016;123(6):1400-1407. [doi: 10.1213/ane.0000000000001362]
7. Ferrante EA, Blasier KR, Givens TB, Lloyd CA, Fischer TJ, Viola F. A novel device for the evaluation of hemostatic function in critical care settings. Anesthesia & Analgesia 2016;123(6):1372-1379. [doi: 10.1213/ane.0000000000001413]
8. Solomon C, Ranucci M, Hochleitner G, Schöchl H, Schlimp CJ. Assessing the methodology for calculating platelet contribution to clot strength (platelet component) in thromboelastometry and thrombelastography. Anesthesia & Analgesia 2015;121(4):868-878. [doi: 10.1213/ane.0000000000000859]

Abbreviations

A: amplitude
ADP: adenosine-diphosphate
G: shear modulus
POCT: point-of-care testing
QG: Quantra group
ROTEM: rotational thromboelastometry
SG: standard group

Edited by G Eysenbach; submitted 29.11.19; peer-reviewed by J Mersmann, S Casu; comments to author 31.12.19; revised version received 04.01.20; accepted 07.01.20; published 21.04.20.

Please cite as:
Michel S, Piekarski F, Fischer JH, Hettler V, Adam EH, Holzer L, Lotz G, Walther T, Zacharowski K, Raimann FJ
Hemostatis Analyzer-Supported Hemotherapy Algorithm in Cardiac Surgery: Protocol for a Randomized Controlled Monocentric Trial
JMIR Res Protoc 2020;9(4):e17206
URL: http://www.researchprotocols.org/2020/4/e17206/
doi:10.2196/17206
PMID:32314972

Protocol

The Scleroderma Patient-Centered Intervention Network Self-Management Program: Protocol for a Randomized Feasibility Trial

Marie-Eve Carrier[1]; Linda Kwakkenbos[2]; Warren R Nielson[3]; Claire Fedoruk[1]; Karen Nielsen[4]; Katherine Milette[1]; Janet Pope[5]; Tracy Frech[6]; Shadi Gholizadeh[7]; Laura Hummers[8]; Sindhu R Johnson[9,10]; Pamela Piotrowski[11]; Lisa Jewett[1]; Jessica Gordon[12]; Lorinda Chung[13]; Dan Bilsker[14,15]; Kimberly A Turner[1]; Julie Cumin[1]; Joep Welling[16]; Catherine Fortune[4]; Catarina Leite[17]; Karen Gottesman[18]; Maureen Sauve[4,19]; Tatiana S Rodríguez-Reyna[20]; Marie Hudson[1,21]; Maggie Larche[22]; Ward van Breda[23]; Maria E Suarez-Almazor[24]; Susan J Bartlett[21]; Vanessa L Malcarne[25]; Maureen D Mayes[26]; Isabelle Boutron[27,28]; Luc Mouthon[27,28]; Fredrick Wigley[8]; Brett D Thombs[1,21,29,30,31,32]; SPIN Investigators[33]

[1]Lady Davis Institute of the Jewish General Hospital, Montreal, QC, Canada
[2]Behavioural Science Institute, Clinical Psychology, Radboud University, Nijmegen, Netherlands
[3]St Joseph's Health Care, London, ON, Canada
[4]Scleroderma Society of Ontario, Hamilton, ON, Canada
[5]University of Western Ontario, London, ON, Canada
[6]University of Utah, Salt Lake City, UT, United States
[7]California School of Professional Psychology/Alliant, Los Angeles, CA, United States
[8]Johns Hopkins University School of Medicine, Baltimore, MD, United States
[9]Toronto Scleroderma Program, Mount Sinai Hospital & Toronto Western Hospital, Toronto, ON, Canada
[10]Division of Neurosurgery, Department of Surgery, University of Toronto, Toronto, ON, Canada
[11]Private practice – Nutrition, Hamilton, ON, Canada
[12]Hospital for Special Surgery, New York, NY, United States
[13]Departments of Pediatrics, Biomedical Data Science, Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States
[14]Simon Fraser University, Burnaby, BC, Canada
[15]University of British Columbia, Vancouver, BC, Canada
[16]NVLE Dutch patient organization for systemic autoimmune diseases, Utrecht, Netherlands
[17]University of Minho, Braga, Portugal
[18]Scleroderma Foundation, Los Angeles, CA, United States
[19]Scleroderma Canada, Hamilton, ON, Canada
[20]Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico
[21]Department of Medicine, McGill University, Montreal, QC, Canada
[22]McMaster University, Hamilton, ON, Canada
[23]Faculty of Behavioural and Movement Sciences, Vrije University, Amsterdam, Netherlands
[24]University of Texas MD Anderson Cancer Center, Houston, TX, United States
[25]San Diego State University, San Diego, CA, United States
[26]University of Texas McGovern School of Medicine, Houston, TX, United States
[27]Université Paris Descartes, Paris, France
[28]Assistance Publique - Hôpitaux de Paris, Paris, France
[29]Department of Psychiatry, McGill University, Montreal, QC, Canada
[30]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada
[31]Department of Psychology, McGill University, Montreal, QC, Canada
[32]Department of Educational and Counselling Psychology, McGill University, Montreal, QC, Canada
[33]See Acknowledgments

Corresponding Author:
Brett D Thombs
Lady Davis Institute of the Jewish General Hospital

4333 Cote Ste-Catherine Road
Montreal, QC, H3T 1E4
Canada
Phone: 1 5143408222 ext 25112
Email: brett.thombs@mcgill.ca

Abstract

Background: Systemic sclerosis (SSc), or scleroderma, is a rare disease that often results in significant disruptions to activities of daily living and can negatively affect physical and psychological well-being. Because there is no known cure, SSc treatment focuses on reducing symptoms and disability and improving health-related quality of life (HRQoL). Self-management programs are known to increase self-efficacy for disease management in many chronic diseases. The Scleroderma Patient-centered Intervention Network (SPIN) developed a Web-based self-management program (SPIN self-management; SPIN-SELF) to increase self-efficacy for disease management and to improve HRQoL for patients with SSc.

Objective: The proposed study aims to assess the feasibility of conducting a full-scale randomized controlled trial (RCT) of the SPIN-SELF program by evaluating the trial implementation processes, required resources and management, scientific aspects, and participant acceptability and usage of the SPIN-SELF program.

Methods: The SPIN-SELF feasibility trial will be conducted via the SPIN Cohort. The SPIN Cohort was developed as a framework for embedded pragmatic trials using the cohort multiple RCT design. In total, 40 English-speaking SPIN Cohort participants with low disease management self-efficacy (Self-Efficacy for Managing Chronic Disease Scale score ≤7), who have indicated interest in using a Web-based self-management program, will be randomized with a 3:2 ratio into the SPIN-SELF program or usual care for 3 months. Feasibility outcomes include trial implementation processes, required resources and management, scientific aspects, and patient acceptability and usage of the SPIN-SELF program.

Results: Enrollment of the 40 participants occurred between July 5, 2019, and July 27, 2019. By November 25, 2019, data collection of trial outcomes was completed. Data analysis is underway, and results are expected to be published in 2020.

Conclusions: The SPIN-SELF program is a self-help tool that may improve disease-management self-efficacy and improve HRQoL in patients with SSc. The SPIN-SELF feasibility trial will ensure that trial methodology is robust, feasible, and consistent with trial participant expectations. The results will guide adjustments that need to be implemented before undertaking a full-scale RCT of the SPIN-SELF program.

International Registered Report Identifier (IRRID): DERR1-10.2196/16799

(JMIR Res Protoc 2020;9(4):e16799) doi:10.2196/16799

KEYWORDS

feasibility studies; scleroderma, systemic; self-management; trial protocols

Introduction

Background

Rare diseases are chronic, disabling, and often life-threatening medical conditions. Individually, rare diseases affect fewer than 1 in 2000 people, but altogether, more than 1 in 15 people (6%-8%) have a rare disease [1,2]. The burden and impact on health-related quality of life (HRQoL) of most rare diseases are high, and in most cases, there is no therapy that cures or modifies the disease itself [3,4]. Psychological, educational, and rehabilitation interventions could contribute to improving function and alleviating distress for patients, and international rare disease plans have emphasized the need for effective disease management tools to complement basic medical care [5,6].

Self-management programs are widely disseminated and are known to increase patient self-efficacy for disease management and HRQoL in many common chronic diseases [7,8]. In rare diseases, patients must address unique challenges that are not part of generic self-management programs [9]. It is difficult, however, to develop, test, and disseminate self-management programs that meet the needs of patients with rare disease [10]. One reason is that the small number of patients makes it difficult to conduct robust clinical trials and effectively disseminate patient tools [10].

Systemic sclerosis (SSc, or scleroderma) is a rare autoimmune disease that affects the skin and internal organs, including the lungs, gastrointestinal tract, and cardiovascular system [11,12]. SSc is notable for the range of problems faced by people living with the disease, including limitations in physical mobility and hand function, pain, fatigue, sleep disturbance, depression, sexual dysfunction, and body image distress from disfiguring changes in appearance [10,13-16]. Self-management is complex in SSc and must address a wide range of concerns, such as skin care, gastrointestinal symptoms, and disfiguring changes in appearance, which are typically not part of generic self-management programs or self-management programs for more common rheumatic diseases [17,18].

One randomized controlled trial (RCT) comparing the effectiveness of an internet self-management program with an educational book was published recently, and two pre-post

intervention studies of SSc self-management programs have been conducted [19-21]. Of the two pre-post studies, one described a mail-delivered program provided to 49 patients, and the other was a pilot study of an internet-based self-management program with 16 patients [19,20]. The small sample sizes and lack of control groups, however, did not allow conclusions to be drawn about effectiveness. Furthermore, the small number of patients involved in these studies underline the difficulty of conducting adequately powered research in a rare disease context. The published RCT compared the effects on self-efficacy of the same internet self-management program as described in the pilot study [20] with those of an educational book in a sample of patients with SSc from the United States (n=267). There were no statistically significant differences between the two groups, but patients were enthusiastic about the Wed-based program and its content [21]. Thus, robust evidence on the effectiveness and availability of accessible self-management tools for patients with SSc remains limited.

The Scleroderma Patient-centered Intervention Network (SPIN) [22], a collaboration of SSc research centers, clinicians, and patient organizations from Canada, the United States of America, Europe, Mexico, and Australia, was created to address this problem [10]. SPIN has assembled a large multinational patient cohort to collect longitudinal data on patient-reported outcomes in SSc and as a framework for embedding RCTs of electronic health interventions on an ongoing basis. To date, more than 2000 patients with SSc from 47 centers have been enrolled in SPIN's Web-based cohort (currently approximately 1800 active participants).

SPIN's Patient Advisory Board has prioritized the need for an SSc-specific disease management program. As such, the SPIN self-management (SPIN-SELF) program was designed by SPIN investigators based on successful self-management programs for more common diseases, informed by data on SSc-specific patient concerns and coping challenges, patient input obtained through a series of focus groups [9], and direct input from patients. The program includes multiple modules that focus on self-efficacy enhancing strategies and provide the knowledge, skills, and confidence essential to managing the physical, emotional, and social consequences of SSc. The SPIN-SELF program utilizes social modeling through educational videos of patients with SSc who describe their own challenges and what they have done to cope with living with SSc, as well as videos of patients and health professionals who teach key self-management techniques [23].

Objective

We will conduct a full-scale RCT to assess the effectiveness of SPIN-SELF on improving self-efficacy and HRQoL. Before this, a feasibility trial of SPIN-SELF is being conducted to ensure the feasibility of the planned trial methodology and that the Web-based intervention is user-friendly and acceptable to trial participants [24-27]. The outcomes of this study will inform and guide adjustments to the Web-based SPIN-SELF program and the trial procedures that may need to be implemented before undertaking the full-scale RCT of SPIN-SELF.

Methods

Design and Setting

The SPIN-SELF feasibility trial is a parallel, two-arm, multicenter RCT that will be conducted via the SPIN Cohort (ClinicalTrials.gov number NCT03914781). The SPIN-SELF feasibility study is not meant for hypothesis testing or effect size estimation, as the sample size is not appropriate to do so, and testing hypotheses about effectiveness is discouraged in the context of feasibility trials [24-27].

The Scleroderma Patient-Centered Intervention Network Cohort Participants

The SPIN Cohort was developed as a framework for embedded pragmatic trials using the cohort multiple RCT (cmRCT) design. In the cmRCT design [28], participants enroll in an observational cohort with regular outcome measurement. Participants consent to (1) allow their data to be used for observational studies; (2) allow their data to be used to assess intervention trial eligibility and, if eligible, be randomized; (3) if eligible and randomized to the intervention arm of the trial, to be contacted and offered access to the intervention, and if eligible and randomized to usual care, to not notify them that they are involved in the trial usual care group and to use their regularly collected cohort data to evaluate trial outcomes. Trial eligibility can be assessed during regular SPIN Cohort assessments, which occur every 3 months, and trial outcomes can be obtained at the subsequent SPIN Cohort assessment 3 months later.

To be eligible for the SPIN Cohort, patients must be classified as having SSc based on 2013 American College of Rheumatology/European League Against Rheumatism criteria [29] confirmed by a SPIN physician; be ≥18 years old; be able to give informed consent; be fluent in English, French, or Spanish; and be able to respond to questionnaires via the internet. The SPIN Cohort is a convenience sample. Eligible SPIN Cohort patients are recruited at SPIN sites [22] during regular medical visits, and written informed consent is obtained. A medical data form is submitted on the Web by the site to enroll participants.

Scleroderma Patient-Centered Intervention Network-Self Feasibility Participants

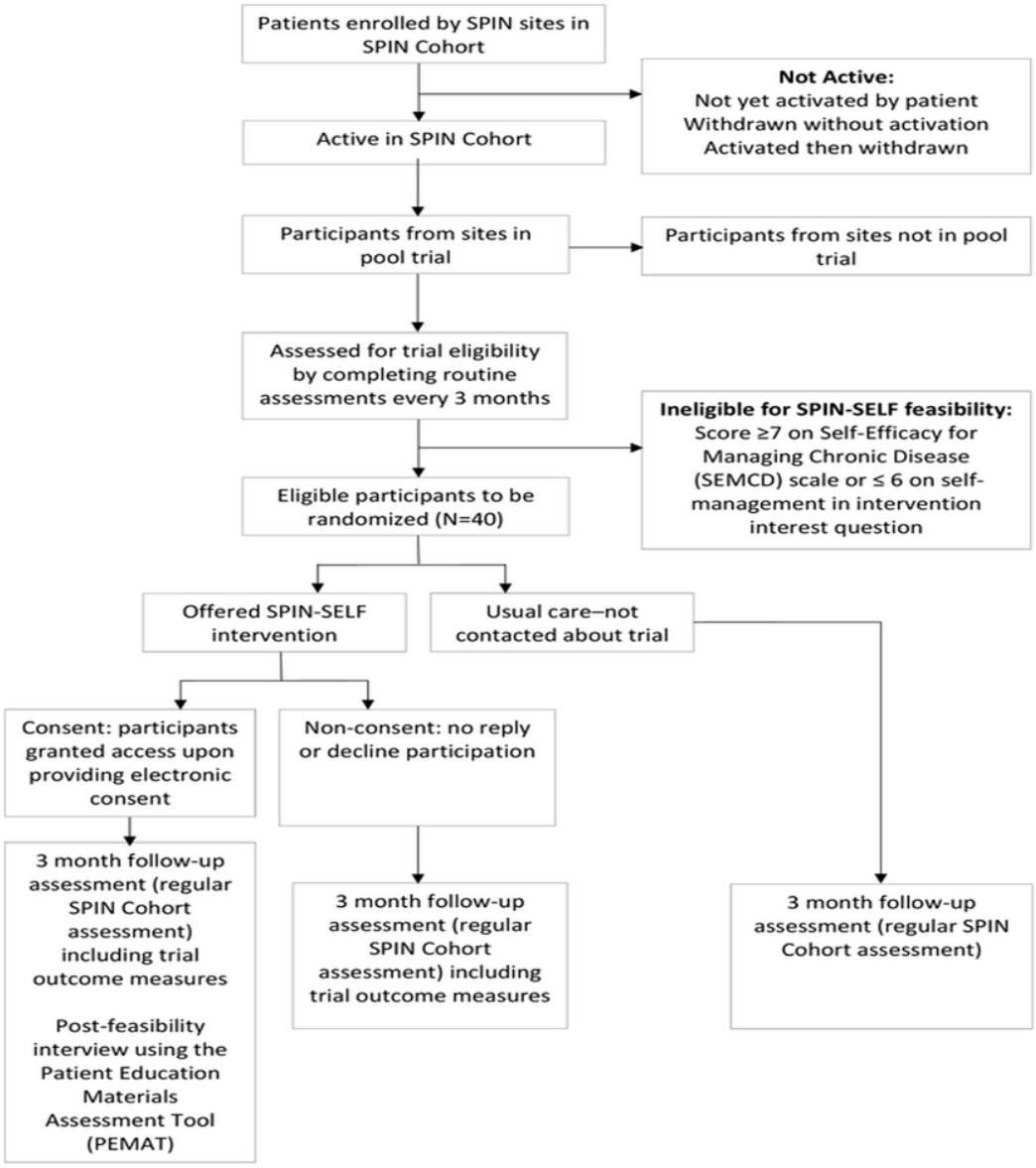
For the SPIN-SELF feasibility trial, 40 English-speaking SPIN Cohort participants will be randomized with a 3:2 ratio to be offered to use the SPIN-SELF program or usual care for 3 months. Cohort participants will be eligible for the feasibility trial if they complete their SPIN Cohort measures in English, have low disease management self-efficacy (Self-Efficacy for Managing Chronic Disease Scale score [SEMCD]≤7 [30]), and have indicated high interest in using a Web-based self-management intervention (≥6 on 0-10 scale). Assessments of disease management self-efficacy and interest will occur as part of participants' regular SPIN Cohort assessments.

Procedure: Randomization, Allocation Concealment, Consent, and Blinding

Randomization to be offered vs not offered to the SPIN-SELF intervention will occur at the time of cohort participants' regular

SPIN Cohort assessments. Eligible cohort participants, based on questionnaire responses, will be randomized automatically as they complete their regular SPIN Cohort assessments using a feature in the SPIN Cohort platform, which provides immediate centralized randomization and thus complete allocation sequence concealment. Participants randomized to be offered the intervention will receive an automated email invitation including a link to the SPIN-SELF program site and the SPIN-SELF feasibility study consent form. At initial log-in, they will be prompted to provide written consent to participate in the SPIN-SELF feasibility trial by verifying agreement with consent elements and providing their email address as the signature. Participants who consent will be automatically redirected to the introduction page of the SPIN-SELF program. Patients who log out before agreeing to the terms of the consent form will return to the consent page upon subsequent log-ins. Participants who accept the offer to use the SPIN-SELF program can use the weblink to enter the secure intervention site. SPIN personnel will also contact participants by phone, usually within 48 hours of sending the invitation email, to describe the study, review the consent form, and answer questions. Within the first 10 days, after the SPIN-SELF invitation email is sent, SPIN personnel will attempt to contact participants, up to a maximum of five times, to offer them more information about the study, to answer any questions they may have, and to help them consent or log in to the program. If a participant is still unreachable, a sixth and last attempt will be conducted by SPIN personnel to complete the call protocol at approximately 20 days postinvite. Email and phone technical support will be available to help participants with the consent process and to access and use the intervention site. See Figure 1 for SPIN Cohort participants' flow through the SPIN-SELF feasibility trial.

Figure 1. Scleroderma Patient-centered Intervention Network Cohort and feasibility trial flow diagram.



In pragmatic trials, participants are typically not blinded to intervention status, and possible biases are accepted as part of the response to being offered an intervention, as may occur in practice [31,32]. Disappointment bias, however, can occur in

conventional trial designs when a participant enrolls in a trial to receive an intervention but is allocated to usual care [28,31]. For this reason, in the cmRCT design [28], participants who are not offered an intervention are not notified that they have not been offered the intervention. This replicates actual practice, where patients are not typically advised about treatments that are not options, and reduces the risk of disappointment bias [28,31,32]. All participants in the SPIN Cohort are aware that SPIN will conduct intervention trials and are routinely asked about potential interest in nine possible interventions as part of their regular cohort assessments but are not informed that any particular intervention may be available unless they are offered to try the intervention. Thus, participants who are offered the intervention are not blind to their status, whereas participants assigned to usual care are blind to their participation in the trial.

Intervention and Comparator

The SPIN-SELF program was designed based on key tenets of behavior change that have been incorporated in successful self-management programs for chronic diseases [17,23,33,34], as well as input from focus groups of patients with SSc and SPIN's Patient Advisory Board. Most self-management programs follow a similar format. They include multiple modules that focus on self-efficacy enhancing strategies and provide the knowledge, skills, and confidence essential for managing the physical, emotional, and social consequences of a disease. Patients are not given direct solutions to problems, but rather are taught problem-solving and management skills. Each module includes an educational component, teaching of skills, and a goal-setting component [17,33,34]. The SPIN-SELF program utilizes social modeling through educational videos of patients with SSc who describe their own challenges and what they have done to cope with living with SSc, as well as videos of patients and health professionals who teach key self-management techniques [23].

After an introduction to self-management by a physician with expertise in the treatment of SSc (Figure 2), a patient shares her experience with learning how to become an efficient self-manager (Figure 3). Instructions on how to navigate the program are provided in a website tour video. Participants are then directed to a nine-item quiz that provides guidance to modules that are most relevant to a patient's symptoms and disease management challenges (Figure 4).

Figure 2. Scleroderma Patient-centered Intervention Network health care provider discussing SPIN-self-management Program.

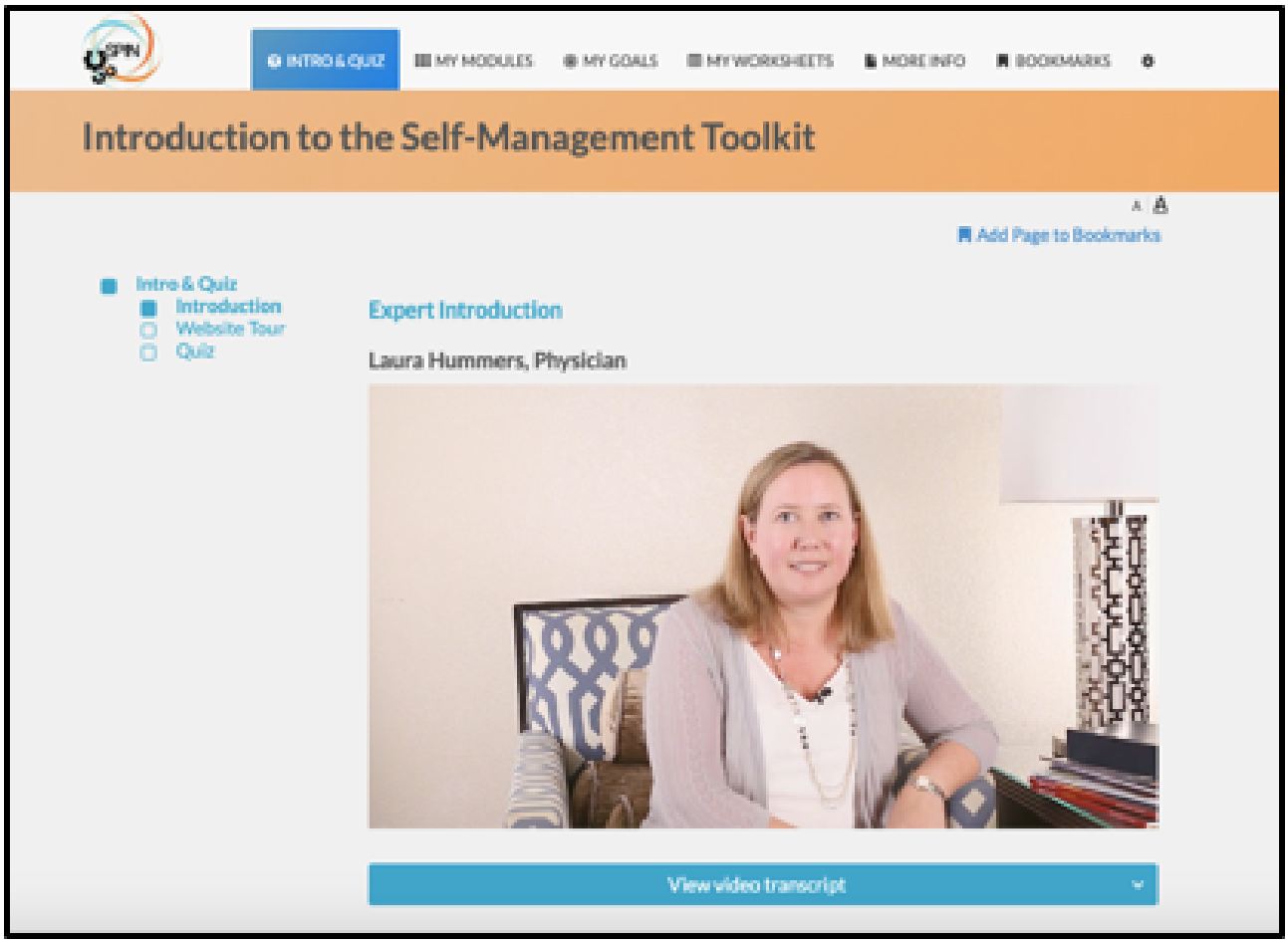


Figure 3. Patient discussing her experience with self-management.

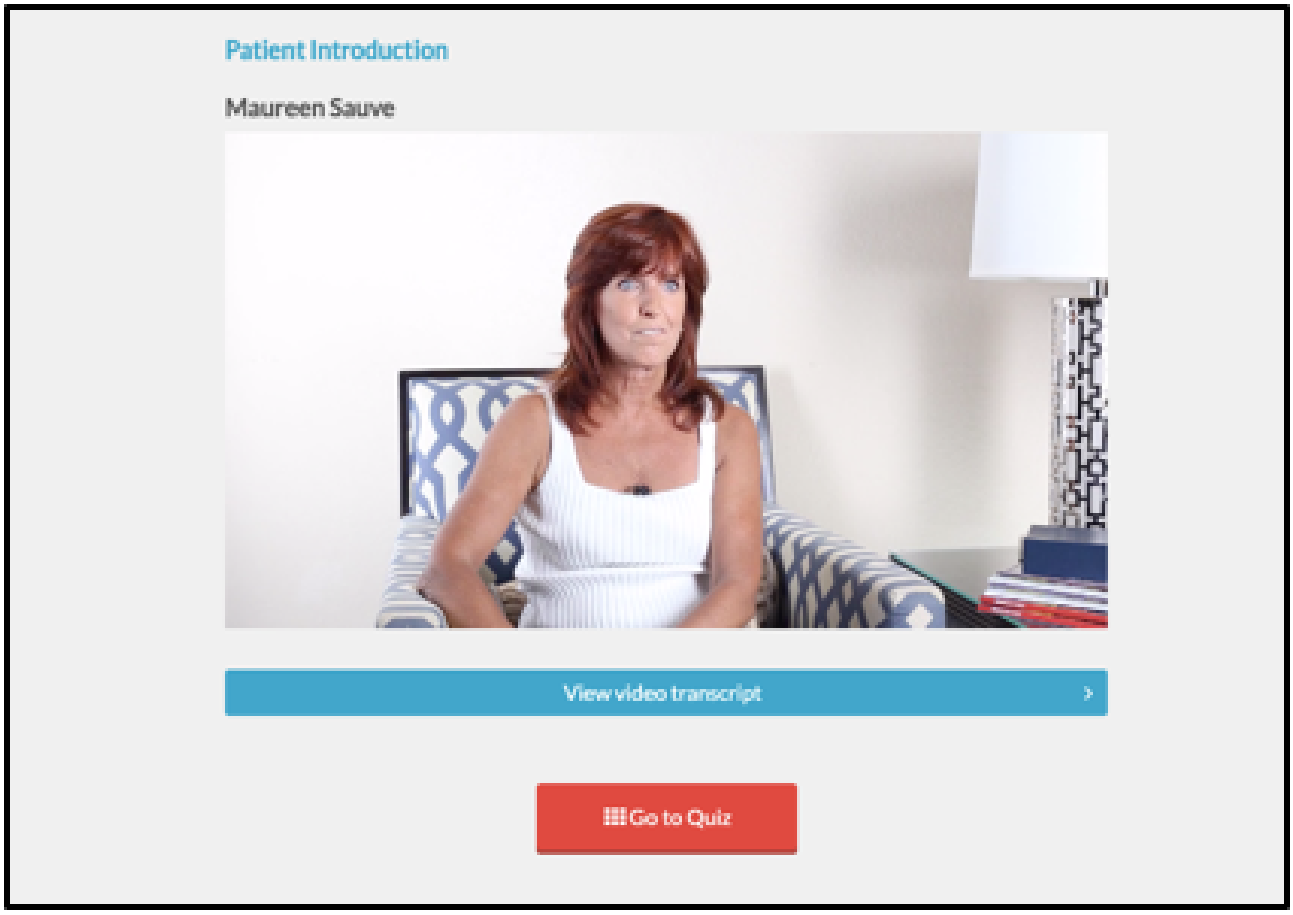
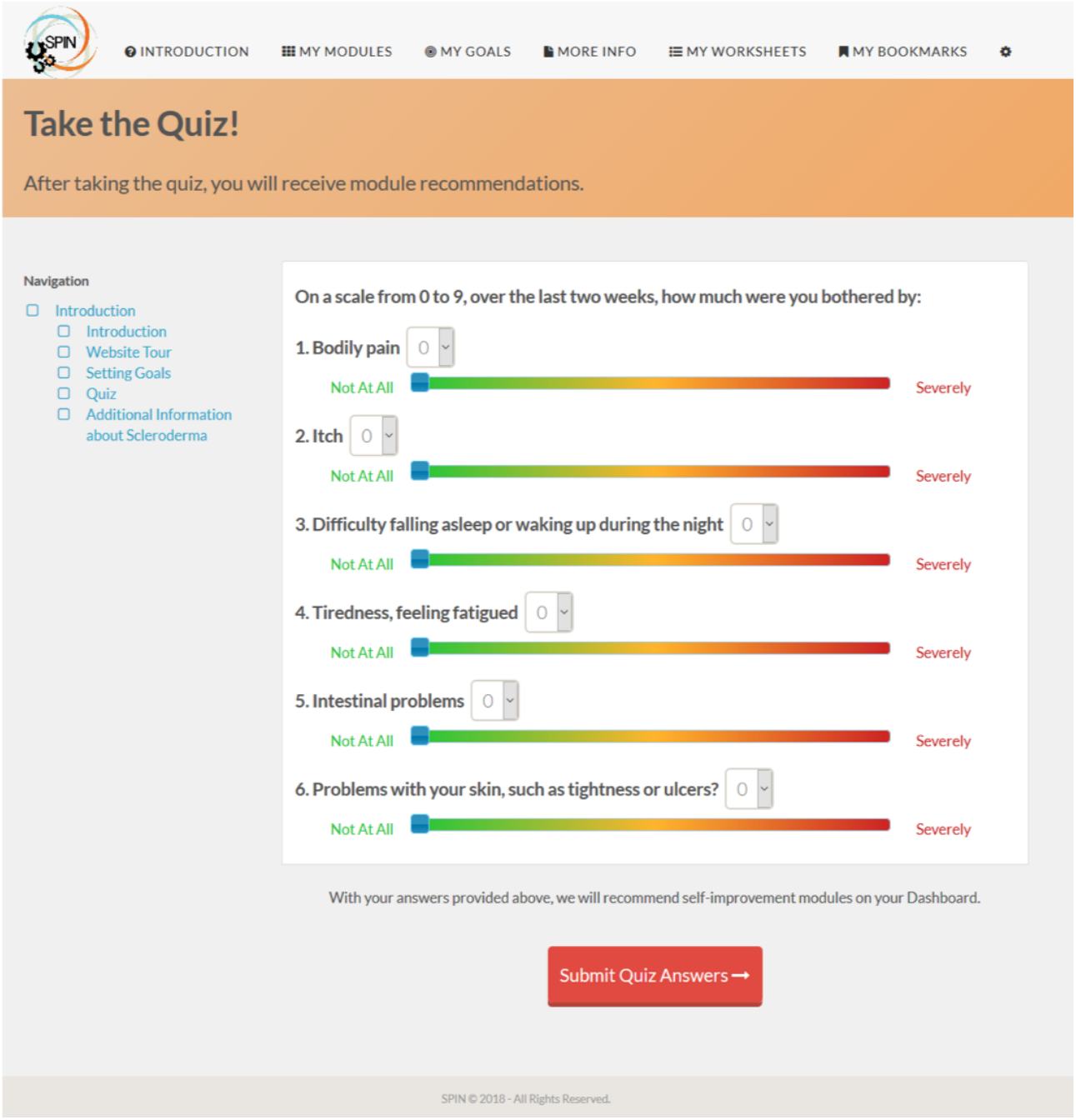


Figure 4. Nine-item quiz provides guidance to modules that are most relevant to a patients' symptoms and disease management challenges.



The program's nine modules focus on (1) coping with pain; (2) skincare, finger ulcers and Raynaud's; (3) sleep problems; (4) fatigue; (5) gastrointestinal symptoms; (6) itch; (7) managing emotions and stress; (8) coping with body image concerns due to disfigurement; and (9) effective communication with health care providers. On the basis of the quiz score, the three modules that are most relevant for the patient will appear on top. Patients may access any modules, including the three identified as most relevant and all other modules; the other modules are shown underneath the three most relevant ones. Access to the nine modules in the program is unrestricted for the duration of the trial (Figure 5).

In addition to core modules, the SPIN-SELF program includes tools to support key components of successful self-management programs, including goal-setting strategies, goal forms, and worksheets to learn how to integrate newly learned skills and techniques into a daily routine [35] (Figure 6). For each goal that patients set, it is possible to input weekly progress, to share goals and progress with friends and family, and receive email reminders (Figure 7). Under the More Info section, general information on SSc can be found, in addition to patient stories of experiences with tips to improve self-management (Figure 8).

The program utilizes an engaging and easy-to-navigate Web interface. Favourite pages can be bookmarked for easy access, and text can be enlarged on every page. Scripts are available for each video.

Figure 5. Menu of Scleroderma Patient-centered Intervention Network self-management modules.

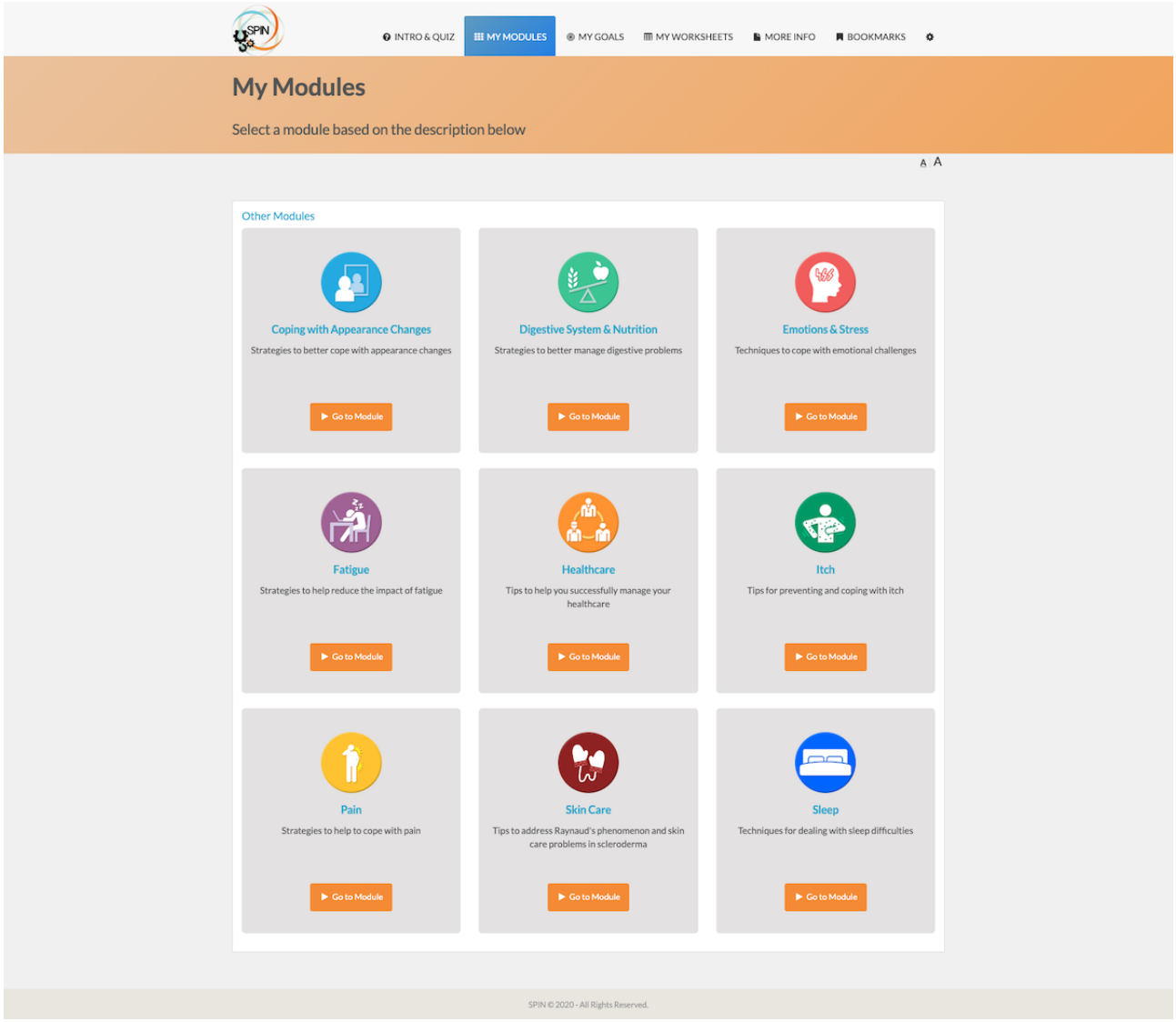


Figure 6. Tools to support key components of successful self-management programs, including goal-setting strategies, goal forms, and worksheets to learn how to integrate newly learned skills and techniques into a daily routine.

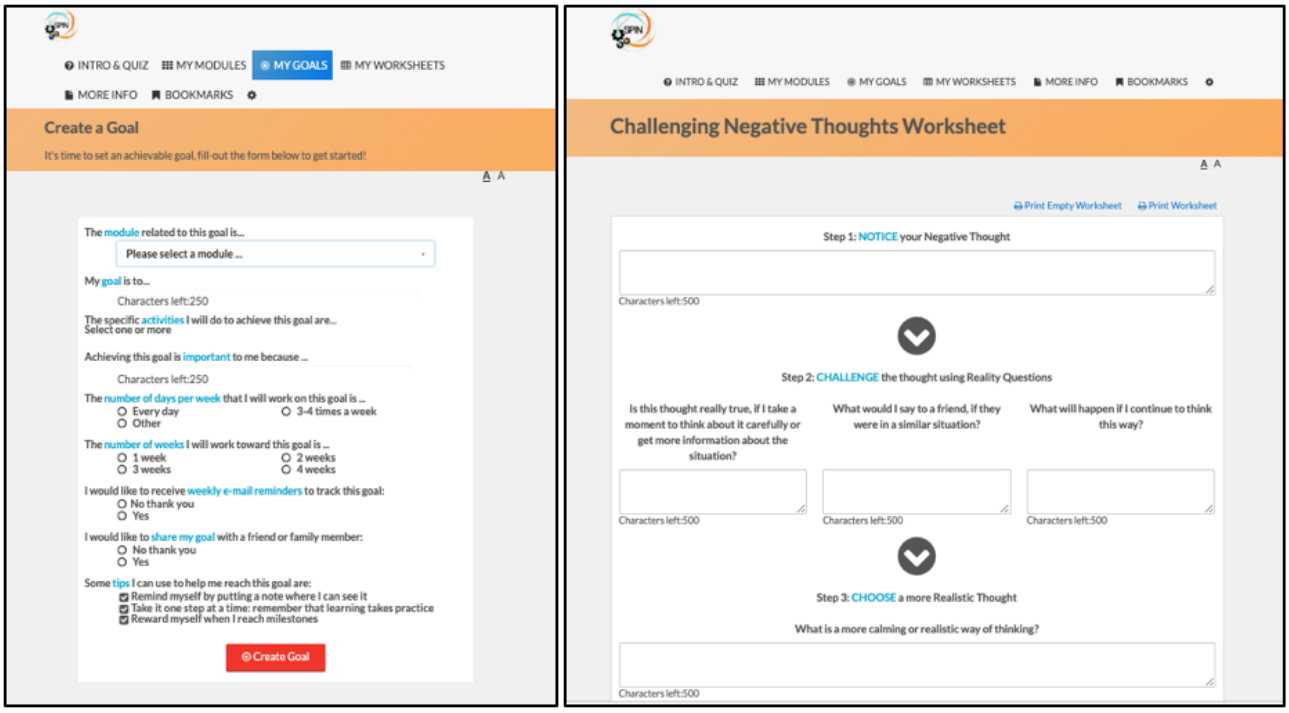


Figure 7. Goals and progress tracking tool.

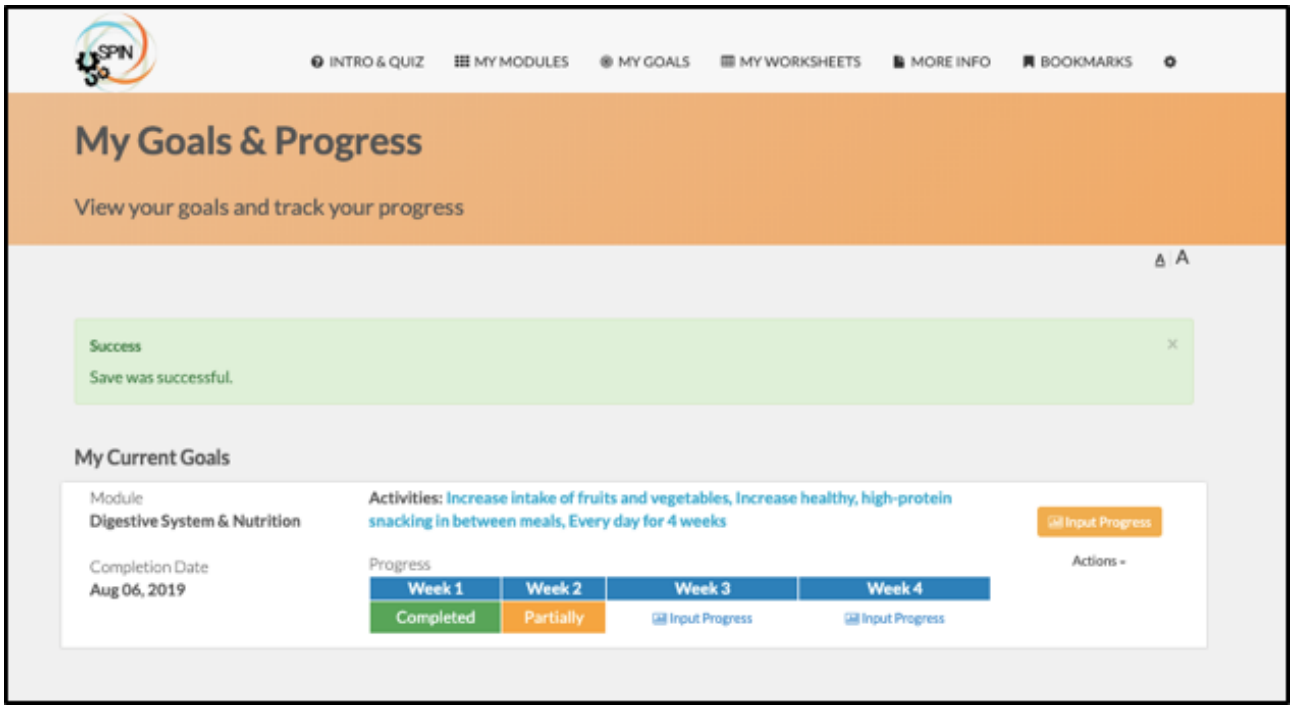
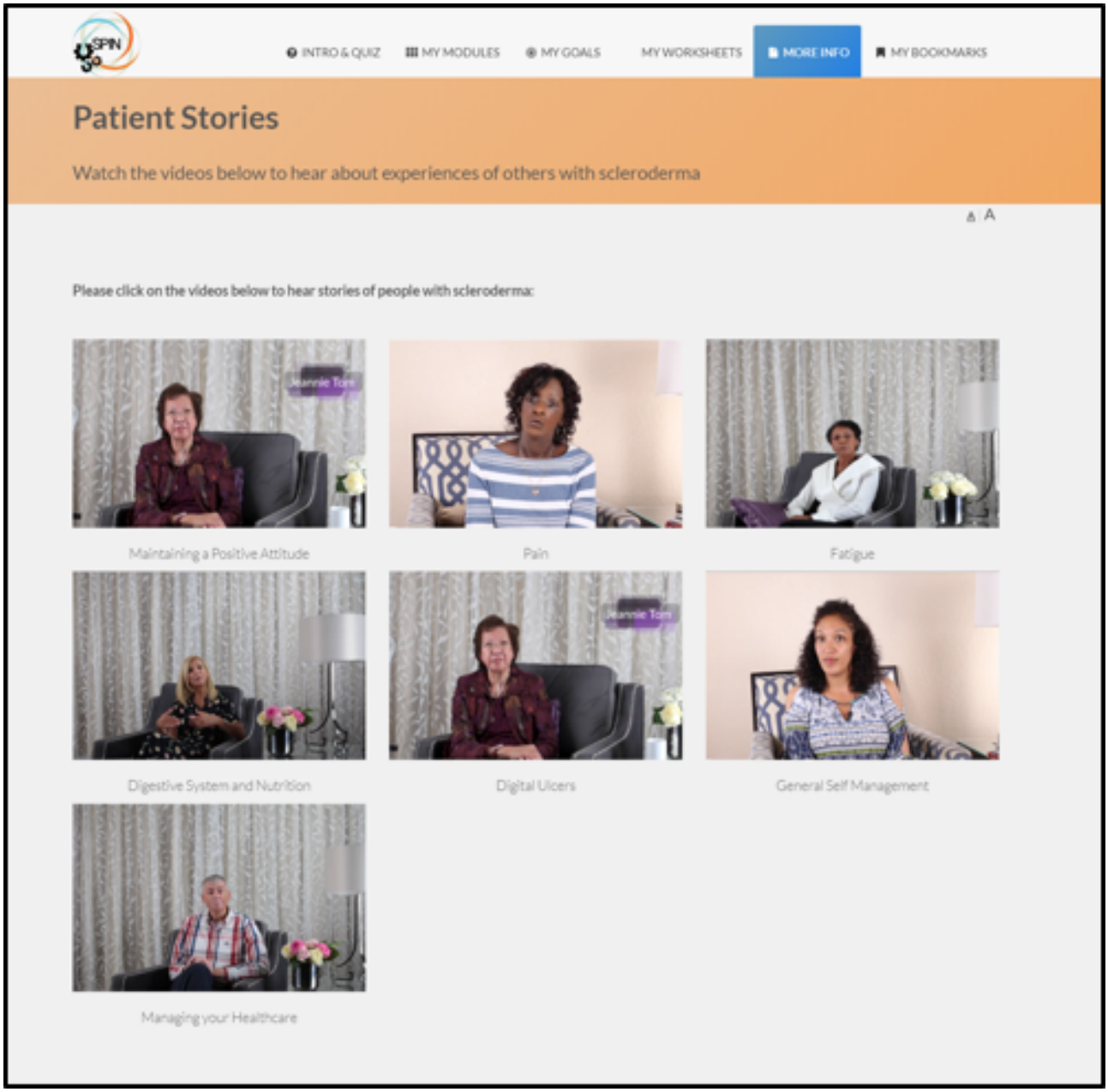


Figure 8. Patient stories page.



Outcomes and Measurement

The primary aim of the SPIN-SELF feasibility trial is to collect data related to the study's *process* to assess the feasibility of the steps that need to take place as part of the full-scale RCT; *required resources and management* (eg, personnel and data management issues); and *scientific aspects* (outcome assessments). Data will be used to determine whether it is feasible to conduct the main study or whether changes need to be made before conducting a full-scale RCT of the SPIN-SELF program. The feasibility trial outcomes related to the process and resources will be assessed throughout the feasibility trial, and patient feedback will be obtained at 3 months postrandomization.

Process and Resources

Information to be collected includes (1) the proportion of SPIN Cohort participants who meet eligibility criteria, (2) proper functioning of automated eligibility and randomization procedures, (3) the proportion of eligible participants randomized to be offered the SPIN-SELF intervention who accept the offer and consent to participate, (4) completeness of Web-based data collection for each trial arm at 3-month follow-up, (5) completeness of the intervention usage log data, (6) ability to successfully link data coming from the SPIN Cohort and SPIN-SELF platforms, (7) rate of completion of trial outcome variables; (8) personnel requirements to call enrolled participants and help them to consent and use the SPIN-SELF program, (9) other challenges for study personnel, and (10) technological performance of the Web-based SPIN-SELF program.

Participant Use and Acceptability of Scleroderma Patient-Centered Intervention Network-Self Program

Usage of the SPIN-SELF program modules among participants in the intervention arm will be examined via intervention usage

data. These data will provide detailed information on the number of log-ins, the number of modules accessed, goals set, as well as time spent on each webpage. In addition, at 3 months postrandomization, qualitative interviews will be conducted with participants in the intervention arm to assess user acceptability and satisfaction. The semistructured interview will be guided by items of the Patient Education Materials Assessment Tool for audiovisual materials [36] and will address topics related to usability, understandability, organization, and clarity. Participant feedback from these interviews will inform any changes necessary to improve the SPIN-SELF program before conducting a full-scale RCT. See Multimedia Appendix 1.

Full-Scale Trial Measures

The objective of the SPIN-SELF full-scale RCT is to evaluate the effect of being offered access to SPIN's Web-based self-management program, in addition to usual care, on disease management self-efficacy and functional health outcomes for patients with SSc with low disease management self-efficacy, compared with usual care alone. In this feasibility study and the full-scale RCT, outcome measures that are routinely assessed as part of the SPIN Cohort assessments every 3 months will be evaluated. For the present feasibility study, we will assess variability in outcome measures and completion rates of outcome variables.

The primary outcome for the full-scale SPIN-SELF trial is *disease management self-efficacy*, which will be evaluated using the SEMCD [30]. The 6-item SEMCD scale measures confidence in the ability to manage fatigue, pain, emotional distress, and other symptoms, to do things to reduce illness impact other than taking medication and to carry out tasks and activities that may reduce the need to see a doctor. Items are rated on a 1 (*not confident at all*) to 10 (*totally confident*) scale. The total score is the mean of all items [30]. The SEMCD has been validated for measuring self-efficacy in patients with SSc through the SPIN Cohort [37]. Mean SEMCD scores of current SPIN Cohort patients who meet SPIN-SELF eligibility criteria is 5.1, which is similar to baseline scores in previous successful trials of self-management programs in other diseases [30].

Patient-reported health status will be measured using the 29-item Patient-Reported Outcomes Measurement Information System (PROMIS-29) profile version 2.0 scale. The PROMIS-29 measures eight domains of health status with 4 items for each of seven domains (physical function, anxiety, depression, fatigue, sleep disturbance, social roles and activities, pain interference) plus a single item for pain intensity. Items are scored on a 5-point scale (range 1-5), with different response options for different domains, and the single pain intensity item is measured on an 11-point rating scale. Higher scores represent more of the domain being measured, that is, better physical function and ability to participate in social roles and activities, but higher levels of anxiety, depression, fatigue, sleep disturbance, pain interference, and pain intensity. Total raw scores are obtained by summing item scores for each domain, which are converted into T-scores standardized from the general US population (mean 50, SD 10). The PROMIS-29 version 2.0 has been validated in SSc[38,39]

Sample Size

Guidance on the appropriate sample size for feasibility trials varies substantially in the published literature, with rules-of-thumb varying from 12 to 30 or more per trial arm [40,41]. To ensure that we collect sufficient quantitative and qualitative outcome data to meet our feasibility objectives and guide the next study phase, we will include a total of 40 SPIN Cohort participants in this feasibility trial. We will randomize using a 3:2 ratio to get a sufficiently large number of patients in the intervention arm. In a previous feasibility trial of a SPIN hand exercises intervention, the rate of acceptance of the offer was approximately 60% [42].

Data Collection, Storage and Sharing

Outcome measures are completed through the participants' regular SPIN Cohort assessments. The SPIN Cohort uses a secure electronic data management platform designed and managed by the *Information Management Services of the Centre for Clinical Epidemiology, Jewish General Hospital, Montreal*. Separate from the SPIN Cohort portal, an encrypted database has been created for the SPIN-SELF program, which includes an identification number for intervention participants to link to SPIN Cohort data and their usage log information.

Data Analysis

The primary data analysis will present a description of feasibility outcomes, including participants' eligibility and recruitment and numbers and percentages of participants who respond to follow-up measures. Use of the internet intervention will be described by presenting the frequency of log-ins and the time spent on the SPIN-SELF program. Analysis of outcome measures will include the completeness of data and the presence of floor or ceiling effects. Descriptive statistics will be used to provide means and SDs for the measures. Qualitative information on participants' experience using the SPIN-SELF intervention will be used to interpret acceptability related to the content, webpage visuals, and navigation and make any necessary changes to the intervention. Information related to the required resources and management of the program during feasibility will inform any necessary changes to intervention or trial procedures.

Data Monitoring

The feasibility trial will be overseen by the SPIN Steering Committee, along with the trial investigators. The Steering Committee will provide scientific direction for the feasibility of the RCT. Routine monitoring of data quality will be handled by the SPIN director, in conjunction with trial investigators. The Steering Committee will be updated on progress during and after the trial.

Adverse Events

The risk of adverse events occurring as a consequence of the SPIN-SELF program is very low. The only risks of participation in the SPIN-SELF feasibility study may be the possible unease or discomfort that some may experience resulting from viewing videos of others living with SSc, from answering questions about their SSc, or from reading the content about SSc in the program. Nonetheless, adverse events will be assessed via

interviews and open-ended questions. Any events reported will be discussed with clinical members of the team, and referrals to local SPIN physicians will be made as necessary. Any serious adverse events that occur will also be reported to the ethics committee.

Ethics and Trial Registration

Ethics approval for the SPIN-SELF feasibility trial has been obtained from the research ethics committee of the Jewish General Hospital, Montreal, Canada. The SPIN-SELF feasibility study was registered before participant enrollment (NCT03914781) and will be reported in accordance with standards articulated in the Consolidated Standards of Reporting Trials (CONSORT) extensions for randomized pilot and feasibility trials [27] and the draft CONSORT extension for trials using cohorts and routinely collected health data, which is forthcoming [43].

Results

Enrollment of the target number of 40 participants occurred between July 5, 2019, and July 27, 2019. By November 25, 2019, data collection of trial outcomes was completed. Data analysis is underway, and results are expected to be published in 2020.

Discussion

The SPIN-SELF program may improve self-efficacy for disease management and HRQoL in patients with SSc. This feasibility study will ensure that trial methodology is robust, feasible, and consistent with participant expectations [24-27]. Results will guide any changes that need to be implemented before conducting a full-scale RCT to test the effectiveness of the SPIN-SELF program. If effective, it will be made available through patient organizations around the world to support people in their efforts to cope with living with SSc.

Acknowledgments

The SPIN-SELF program was made possible, thanks to the hard work and dedication of the SPIN Advisory Board members including Catherine Fortune, Scleroderma Society of Ontario, Hamilton, Ontario, Canada; Dominique Godard, Association des Sclérodermiques de France, Sorel-Moussel, France; Karen Gottesman, Scleroderma Foundation, Los Angeles, California, United States; Geneviève Guillot, Sclérodermie Québec, Montreal, Quebec, Canada; Catarina Leite, University of Minho, Braga, Portugal; Karen Nielson, Scleroderma Society of Ontario, Hamilton, Ontario, Canada; Alexandra Portales, Asociación Española de Esclerodermia, Madrid, Spain; Maureen Sauve, Scleroderma Society of Ontario, Hamilton, Ontario, Canada; and Joep Welling, NVLE Dutch patient organization for systemic autoimmune diseases, Utrecht, The Netherlands. Funding for the SPIN-SELF Feasibility trial was provided by the Canadian Institutes of Health Research (TR3-119192; PJT-148504). SPIN has also been funded by grants from the Arthritis Society and The Canadian Institute for Outcomes in Rheumatology Care. In addition, SPIN has received contributions from the Lady Davis Institute for Medical Research of the Jewish General Hospital, Montreal, Canada, and from McGill University, Montreal, Canada. SPIN has also received support from the Scleroderma Society of Ontario, Scleroderma Canada, Sclérodermie Québec, Scleroderma Manitoba, the Scleroderma Society of Nova Scotia, the Scleroderma Association of British Columbia, and the Scleroderma Association of Saskatchewan. LK was supported by a Canadian Institutes of Health Research Banting Postdoctoral Fellowship. BT was supported by a Fonds de recherche du Québec-Santé researcher salary award. Funders of the SPIN-SELF feasibility trial had no role in the study design or writing and publishing decisions.

Authors' Contributions

BT, LK, MC, SB, VM, MM, IB, LM, FW, WN, KG, and MS were responsible for the study conception and design. BT, LK, MC, WN, CFe, KN, KM, JP, TF, SG, LH, SJ, PP, LJ, JG, LC, DB, KT, JC, JW, CFo, KG, MS, TR, MH, ML, MA, SB, VM, MM, LM, and FW contributed to content development. BT, LK, IB, MC, CFe, KT, and JC will be responsible for the implementation of the trial or acquisition, analysis, and interpretation of trial data. MC drafted the protocol manuscript. All authors provided a critical review and approved the final manuscript. BT is the guarantor.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide assessing usability, understandability, organization, and clarity of the Scleroderma Patient-centered Intervention Network self-management program.
[DOCX File , 15 KB - resprot_v9i4e16799_app1.docx]

Multimedia Appendix 2
Peer-reviewer report by Canadian Institutes of Health Research.
[PDF File (Adobe PDF File), 479 KB - resprot_v9i4e16799_app2.pdf]

References

1. Eurordis Rare Disease Europe. URL: https://www.eurordis.org/about-rare-diseases [accessed 2020-01-11]
2. Rare Disease Foundation. URL: https://www.rarediseasefoundation.org [accessed 2020-01-11]
3. Schieppati A, Henter J, Daina E, Aperia A. Why rare diseases are an important medical and social issue. Lancet 2008 Jun 14;371(9629):2039-2041. [doi: 10.1016/S0140-6736(08)60872-7] [Medline: 18555915]
4. Kole A, Faurisson F. EURORDIS. The Voice of 12,000 Patients: Experiences and Expectations of Rare Disease Patients on Diagnosis and Care in Europe URL: https://www.eurordis.org/IMG/pdf/voice_12000_patients/EURORDISCARE_FULLBOOKr.pdf [accessed 2020-01-11]
5. Canadian Organization for Rare Disorders. 2015 May. Now is the Time: A Strategy for Rare Diseases in a Strategy for All Canadians URL: https://www.raredisorders.ca/content/uploads/CORD_Canada_RD_Strategy_22May15.pdf [accessed 202-01-11]
6. European Commission. National plans or strategies for rare diseases URL: https://ec.europa.eu/health/rare_diseases/national_plans/detailed/index_en.htm [accessed 2020-01-11]
7. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]
8. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008 Jul 15;59(7):1009-1017 [FREE Full text] [doi: 10.1002/art.23817] [Medline: 18576310]
9. Milette K, Thombs BD, Maiorino K, Nielson WR, Körner A, Peláez S. Challenges and strategies for coping with scleroderma: implications for a scleroderma-specific self-management program. Disabil Rehabil 2019 Oct;41(21):2506-2515. [doi: 10.1080/09638288.2018.1470263] [Medline: 29741963]
10. Kwakkenbos L, Jewett LR, Baron M, Bartlett SJ, Furst D, Gottesman K, et al. The Scleroderma Patient-centered Intervention Network (SPIN) Cohort: protocol for a cohort multiple randomised controlled trial (cmRCT) design to support trials of psychosocial and rehabilitation interventions in a rare disease context. BMJ Open 2013 Aug 7;3(8):pii: e003563 [FREE Full text] [doi: 10.1136/bmjopen-2013-003563] [Medline: 23929922]
11. Kowal-Bielecka O, Landewé R, Avouac J, Chwiesko S, Miniati I, Czirjak L, EUSTAR Co-Authors. EULAR recommendations for the treatment of systemic sclerosis: a report from the EULAR Scleroderma Trials and Research group (EUSTAR). Ann Rheum Dis 2009 May;68(5):620-628. [doi: 10.1136/ard.2008.096677] [Medline: 19147617]
12. Seibold J. Scleroderma. In: Harris ED, Budd RC, Firestein GS, editors. Kelley's Textbook of Rheumatology. Seventh Edition. Philadelphia, PA: Elsevier; 2005:1279-1308.
13. Wigley FM, Hummers LK. Clinical features of systemic sclerosis. In: Hochberg MC, Silman AJ, Smolen JS, Weinblatt ME, Weisman MH, editors. Rheumatology. Philadelphia: Mosby; 2003:1463-1480.
14. Bassel M, Hudson M, Taillefer SS, Schieir O, Baron M, Thombs BD. Frequency and impact of symptoms experienced by patients with systemic sclerosis: results from a Canadian National Survey. Rheumatology (Oxford) 2011 Apr;50(4):762-767. [doi: 10.1093/rheumatology/keq310] [Medline: 21149249]
15. Thombs BD, van Lankveld W, Bassel M, Baron M, Buzza R, Haslam S, et al. Psychological health and well-being in systemic sclerosis: State of the science and consensus research agenda. Arthritis Care Res (Hoboken) 2010 Aug;62(8):1181-1189 [FREE Full text] [doi: 10.1002/acr.20187] [Medline: 20235217]
16. Kwakkenbos L, Delisle VC, Fox RS, Gholizadeh S, Jewett LR, Levis B, et al. Psychosocial aspects of scleroderma. Rheum Dis Clin North Am 2015 Aug;41(3):519-528. [doi: 10.1016/j.rdc.2015.04.010] [Medline: 26210133]
17. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
18. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262. [Medline: 11769298]
19. Poole JL, Skipper B, Mendelson C. Evaluation of a mail-delivered, print-format, self-management program for persons with systemic sclerosis. Clin Rheumatol 2013 Sep;32(9):1393-1398. [doi: 10.1007/s10067-013-2282-7] [Medline: 23652719]
20. Poole JL, Mendelson C, Skipper B, Khanna D. Taking charge of systemic sclerosis: a pilot study to assess the effectiveness of an internet self-management program. Arthritis Care Res (Hoboken) 2014 May;66(5):778-782 [FREE Full text] [doi: 10.1002/acr.22192] [Medline: 24115761]
21. Khanna D, Serrano J, Berrocal VJ, Silver RM, Cuencas P, Newbill SL, et al. Randomized controlled trial to evaluate an internet-based self-management program in systemic sclerosis. Arthritis Care Res (Hoboken) 2019 Mar;71(3):435-447. [doi: 10.1002/acr.23595] [Medline: 29741230]
22. The Scleroderma Patient-centered Intervention Network. URL: http://www.spinsclero.com [accessed 2020-01-11]
23. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
24. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry 2006 May;63(5):484-489. [doi: 10.1001/archpsyc.63.5.484] [Medline: 16651505]

25. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
26. van Teijlingen ER, Rennie A, Hundley V, Graham W. The importance of conducting and reporting pilot studies: the example of the Scottish Births Survey. J Adv Nurs 2001 May;34(3):289-295. [doi: 10.1046/j.1365-2648.2001.01757.x] [Medline: 11328433]
27. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Br Med J 2016 Oct 24;355:i5239 [FREE Full text] [doi: 10.1136/bmj.i5239] [Medline: 27777223]
28. Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the 'cohort multiple randomised controlled trial' design. Br Med J 2010 Mar 19;340:c1066. [doi: 10.1136/bmj.c1066] [Medline: 20304934]
29. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum 2013 Nov;65(11):2737-2747 [FREE Full text] [doi: 10.1002/art.38098] [Medline: 24122180]
30. Ritter PL, Lorig K. The English and Spanish Self-Efficacy to Manage Chronic Disease Scale measures were validated using multiple studies. J Clin Epidemiol 2014 Nov;67(11):1265-1273. [doi: 10.1016/j.jclinepi.2014.06.009] [Medline: 25091546]
31. Torgerson DJ, Torgerson CJ. Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction. London, United Kingdom: Palgrave Macmillan; 2008.
32. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, CONSORT group, Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. Br Med J 2008 Nov 11;337:a2390 [FREE Full text] [doi: 10.1136/bmj.a2390] [Medline: 19001484]
33. Marks R, Allegrante JP, Lorig K. A review and synthesis of research evidence for self-efficacy-enhancing interventions for reducing chronic disability: implications for health education practice (part I). Health Promot Pract 2005 Jan;6(1):37-43. [doi: 10.1177/1524839904266790] [Medline: 15574526]
34. Marks R, Allegrante JP, Lorig K. A review and synthesis of research evidence for self-efficacy-enhancing interventions for reducing chronic disability: implications for health education practice (part II). Health Promot Pract 2005 Apr;6(2):148-156. [doi: 10.1177/1524839904266792] [Medline: 15855284]
35. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/s0738-3991(02)00032-0] [Medline: 12401421]
36. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. Patient Educ Couns 2014 Sep;96(3):395-403 [FREE Full text] [doi: 10.1016/j.pec.2014.05.027] [Medline: 24973195]
37. Riehm KE, Kwakkenbos L, Carrier M, Bartlett SJ, Malcarne VL, Mouthon L, Scleroderma Patient-Centered Intervention Network Investigators. Validation of the self-efficacy for managing chronic disease scale: a scleroderma patient-centered intervention network cohort study. Arthritis Care Res (Hoboken) 2016 Aug;68(8):1195-1200 [FREE Full text] [doi: 10.1002/acr.22807] [Medline: 26619042]
38. Hinchcliff M, Beaumont JL, Thavarajah K, Varga J, Chung A, Podlusky S, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-Reported Outcomes Measurement Information System 29-item Health Profile and Functional Assessment of Chronic Illness Therapy-Dyspnea short form. Arthritis Care Res (Hoboken) 2011 Nov;63(11):1620-1628 [FREE Full text] [doi: 10.1002/acr.20591] [Medline: 22034123]
39. Kwakkenbos L, Thombs B, Khanna D, Carrier M, Baron M, Furst D, SPIN Investigators. Performance of the Patient-Reported Outcomes Measurement Information System-29 in scleroderma: a Scleroderma Patient-centered Intervention Network Cohort Study. Rheumatology (Oxford) 2017 Aug 1;56(8):1302-1311 [FREE Full text] [doi: 10.1093/rheumatology/kex055] [Medline: 28431140]
40. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005;4(4):287-291. [doi: 10.1002/pst.185]
41. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]
42. Kwakkenbos L, Carrier M, Welling J. Feasibility Trial of an Internet-based Exercise Program to Improve Hand Function in Patients With Scleroderma: A Scleroderma Patient-centered Intervention Network (SPIN) Study. ClinicalTrials.gov ID:NCT03092024. In progress. URL: https://clinicaltrials.gov/ct2/show/NCT03092024
43. Kwakkenbos L, Juszczak E, Hemkens LG, Sampson M, Fröbert O, Relton C, et al. Protocol for the development of a CONSORT extension for RCTs using cohorts and routinely collected health data. Res Integr Peer Rev 2018;3:9 [FREE Full text] [doi: 10.1186/s41073-018-0053-3] [Medline: 30397513]

Abbreviations

cmRCT: cohort multiple RCT
CONSORT: Consolidated Standards of Reporting Trials
HRQoL: health-related quality of life

PROMIS-29: Patient-Reported Outcomes Measurement Information System
RCT: randomized controlled trial
SEMCD: Self-Efficacy for Managing Chronic Disease Scale score
SPIN: Scleroderma Patient-centered Intervention Network
SPIN-SELF: SPIN self-management
SSc: systemic sclerosis

Edited by G Eysenbach; submitted 13.11.19; peer-reviewed by R Ottwell, P Hsiao; comments to author 11.01.20; revised version received 18.01.20; accepted 24.01.20; published 24.04.20.

Please cite as:
Carrier ME, Kwakkenbos L, Nielson WR, Fedoruk C, Nielsen K, Milette K, Pope J, Frech T, Gholizadeh S, Hummers L, Johnson SR, Piotrowski P, Jewett L, Gordon J, Chung L, Bilsker D, Turner KA, Cumin J, Welling J, Fortune C, Leite C, Gottesman K, Sauve M, Rodríguez-Reyna TS, Hudson M, Larche M, van Breda W, Suarez-Almazor ME, Bartlett SJ, Malcarne VL, Mayes MD, Boutron I, Mouthon L, Wigley F, Thombs BD, SPIN Investigators
The Scleroderma Patient-Centered Intervention Network Self-Management Program: Protocol for a Randomized Feasibility Trial
JMIR Res Protoc 2020;9(4):e16799
URL: http://www.researchprotocols.org/2020/4/e16799/
doi:10.2196/16799
PMID:32329747

Protocol

Implementation and Effects of Information Technology-Based and Print-Based Interventions to Promote Physical Activity Among Community-Dwelling Older Adults: Protocol for a Randomized Crossover Trial

Claudia R Pischke[1*], PhD; Claudia Voelcker-Rehage[2*], PhD; Manuela Peters[3*], MSc; Tiara Ratz[4*], MSc; Hermann Pohlabeln[3*], PhD; Jochen Meyer[5*], PhD; Kai von Holdt[5*], BSc; Sonia Lippke[4*], PhD

[1]Institute of Medical Sociology, Centre for Health and Society, Medical Faculty, University of Duesseldorf, Duesseldorf, Germany

[2]Institute of Human Movement Science and Health, Chemnitz University of Technology, Chemnitz, Germany

[3]Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

[4]Jacobs University Bremen, Bremen, Germany

[5]OFFIS – Institute for Information Technology, Oldenburg, Germany

[*]all authors contributed equally

Corresponding Author:
Claudia R Pischke, PhD
Institute of Medical Sociology
Centre for Health and Society
Medical Faculty, University of Duesseldorf
Moorenstrasse 5
Duesseldorf, 40225
Germany
Phone: 49 49 211 81 ext 08599
Email: claudiaruth.pischke@med.uni-duesseldorf.de

Abstract

Background: Despite the known health benefits of physical activity (PA), less than half and less than one-third of older adults in Germany reach the PA recommendations for endurance training and strength training, respectively, of the World Health Organization. The aim of this study is to investigate the implementation and effectiveness over the course of 9 months of two interventions (information technology [IT]-based vs print-based) for PA promotion among initially inactive older adults in a randomized, crossover trial. This study is part of a large research consortium (2015-2021) investigating different aspects of PA promotion. The IT-based intervention was previously developed and refined, while the print-based intervention was newly developed during this funding phase.

Objective: We aim to compare the effectiveness and examine the preferences of study participants regarding both delivery modes.

Methods: Our target sample size was 390 initially inactive community-dwelling older adults aged ≥60 years at baseline (3-month follow-up [T1]: expected n=300; 9-month follow-up [T2]: expected n=240) who were randomized to one of two interventions for self-monitoring PA: IT-based (50%) or print-based (50%) intervention. In addition, 30% of the IT-based intervention group received a PA tracker. At T1, participants in both groups could choose whether they prefered to keep their assigned intervention or cross over to the other group for the following 6 months (T2). Participants' intervention preferences at baseline were collected retrospectively to run a post hoc matched-mismatched analysis. During the initial 3-month intervention period, both intervention groups were offered weekly group sessions that were continued monthly between T1 and T2. A self-administered questionnaire and 3D accelerometers were employed to assess changes in PA between baseline, T1, and T2. Adherence to PA recommendations, attendance at group sessions, and acceptance of the interventions were also tracked.

Results: The funding period started in February 2018 and ends in January 2021. We obtained institutional review board approval for the study from the Medical Association in Bremen on July 3, 2018. Data collection was completed on January 31, 2020, and data cleaning and analysis started in February 2020. We expect to publish the first results by the end of the funding period.

Conclusions: Strategies to promote active aging are of particular relevance in Germany, as 29% of the population is projected to be ≥65 years old by 2030. Regular PA is a key contributor to healthy aging. This study will provide insights into the acceptance and effectiveness of IT-based vs print-based interventions to promote PA in initially inactive individuals aged ≥60 years. Results obtained in this study will improve the existing evidence base on the effectiveness of community-based PA interventions in Germany and will inform efforts to anchor evidence-based PA interventions in community structures and organizations via an allocation of permanent health insurance funds.

Trial Registration: German Registry of Clinical Trials DRKS00016073; https://tinyurl.com/y983586m

International Registered Report Identifier (IRRID): DERR1-10.2196/15168

(JMIR Res Protoc 2020;9(4):e15168) doi:10.2196/15168

KEYWORDS

physical activity; older adults; eHealth; print intervention; IT-based intervention; physical activity promotion; healthy aging; preferences; randomized trial

Introduction

In Germany, approximately 91% of all deaths are attributable to noncommunicable diseases [1]. Physical inactivity is the fourth leading risk factor contributing to the development of noncommunicable diseases and overall mortality [2]. Conversely, regular physical activity (PA) and the reduction of an inactive lifestyle [3] are associated with improvements in physical, cognitive, and functional health over the lifespan [3,4].

The World Health Organization (WHO) and American College of Sports Medicine recommend that adults aged 18 to 64 years, as well as those 65 years and older, should perform moderate-to-vigorous endurance training for at least 150 minutes per week (in bouts of at least 10 minutes) [2]. In addition, adults aged 65 years and older should perform flexibility, strength, and balance training twice per week [5,6]. Furthermore, Rütten and Pfeiffer [7] state in the German national recommendations for PA that adults should generally avoid extended periods of sitting and, where possible, take PA breaks from sitting. According to the authors, “a major health benefit of PA can already be observed when persons who were entirely physically inactive start becoming active to a small extent meaning that every increase in PA is associated with a health benefit. Every single step away from sedentary behavior is important and promotes health” [7].

In Germany, only 42% of adults aged 65 years and older currently meet the recommendations for endurance (at least 150 minutes per week of moderate activity, at least 75 minutes of intense activity, or a combination of both, in bouts of at least 10 minutes), and 29% meet the recommendations for strength training [6], indicating a need for interventions for PA promotion targeting the general population of older adults. An earlier survey spanning the years 2008 to 2011 reported that merely 18% of German adults aged 60 to 69 years and 14% of adults aged 70 to 79 years exercised moderately to vigorously for at least 150 minutes per week [8]. To tackle this public health issue in Germany and facilitate PA promotion, the Federal Ministry of Health recommends the development and implementation of population-based informational intervention approaches or campaigns, community-based interventions, and policy and environmental approaches for PA promotion [7]. According to the current state of research, results of several reviews suggest that interventions for PA promotion, including mass media campaigns, motivational decision aids, community-based multicomponent interventions, and environmental approaches, can effectively increase PA in the general population [7]. Evidence regarding the benefit of using theory during intervention development to achieve greater impact on behavior change is still contradictory [9]. On the one hand, there is some evidence suggesting that interventions based on behavior change techniques rooted in theory are effective in altering behavior [9]. On the other hand, results of one meta-analysis indicate that a theoretical basis leads to no difference in intervention effectiveness [10], maybe also due to the fact that the behavior preferences of study participants were not sufficiently taken into account [11]. Theories and theoretical assumptions about mechanisms, processes, and techniques are needed to better understand individuals and their needs and to avoid always reinventing the wheel when developing interventions but rather building on previous evidence [3,9,10].

Multiple studies have investigated the role of different modalities of delivering these interventions to older adults. Evidence suggests that participation in interventions providing information on PA either in a face-to-face setting or as print versions leads to increased PA levels in older adults [12-14]. Information technology (IT)-based PA interventions have the advantage of potentially reaching a large number of people in a cost-effective manner [15]. In addition, intervention material can be easily accessed, and instantaneous feedback on behavior change can be provided [10,15]. They also appear to have a positive impact on PA [3,16]. Results of a systematic review comparing the effectiveness of electronic health (eHealth) interventions promoting PA in older adults aged 55 years and older with either no intervention or a non-eHealth intervention indicate that eHealth interventions can effectively promote PA in this population in the short-term, while evidence regarding long-term effects is still lacking [3,9,17]. The results of this review are inconclusive regarding the question of whether eHealth interventions have a greater impact on PA behavior among older adults than non-eHealth (eg, print-based) interventions [17]. Also, the effects of combining various eHealth intervention components are still unclear.

Two studies investigated the added benefit of using Fitbit activity trackers in addition to a website to monitor PA on total weekly PA levels and moderate-to-vigorous PA (MVPA)

[18,19]. In a sample of overweight adults, Vandelanotte et al [18] demonstrated a significant increase in total weekly PA and MVPA after 3 months for a group of participants using a Fitbit compared with a group of participants not using a Fitbit. Although sitting time decreased over time, this effect was not significant [18]. With a sample of older adults, Muellmann et al [19] found no significant differences in PA levels between a group using a Fitbit compared with a group not using a Fitbit. However, participants in the Fitbit group had slightly greater increases in MVPA and decreases in sedentary time than participants in the non-Fitbit group after 3 months [19]. While more participants in the non-Fitbit group did not complete the follow-up after 3 months than in the Fitbit group (63% vs 36%) in the first study [18], the reverse was true in the second study [19]. More participants in the Fitbit group did not complete the 3-month follow-up than in the non-Fitbit group (40% vs 31%) [19]. It is conceivable that individual preferences for interventions may have influenced dropout and possibly adherence to the interventions [10,11,13]. Participants may have been randomized to an intervention that they would not have picked had they been given a choice and that they may have found difficult to interact with.

The effects of study participants' preferences on treatment or intervention outcomes in randomized controlled trials are still not well understood [20,21]. There is some indication that the preferences for delivery mode vary by the sociodemographic characteristics of participants, such as age, gender, and living environment, or by weight status [13,22]. For example, preference for an IT-based intervention was positively related with being 35 years or older (compared to younger ages) and high levels of internet use and was negatively associated with female gender. Preference for a print-based intervention was associated with older age and negatively associated with female gender and obesity [22]. Further, evidence suggests that sociodemographic variables may explain variations in the use of PA trackers [23] and that use of trackers in PA interventions should be aligned with preferences of different target groups [13].

Therefore, the main aim of this study was to compare the effects of 2 interventions using different modalities (IT-based vs print-based) among initially inactive older adults aged at least 60 years living in 14 community districts in 2 geographically different regions (northwest and northeast) of the city of Bremen, Germany. Further, participants' preferences regarding the intervention modality were addressed by conducting a randomized trial with a crossover design. At baseline, participants were randomized to either a 10-week IT-based or print-based intervention. A random subsample of the IT-based intervention group (30%) also received a PA tracker. After the 3-month follow-up, participants could choose whether they wanted to remain in the same intervention group and keep their assigned material or switch to the other group for the following 6 months. They could also choose to use a PA tracker at this stage.

Methods

The study is embedded in the larger Physical Activity and Health Equity: Primary Prevention for Healthy Ageing (AEQUIPA) research network that is funded by the Federal Ministry of Education and Research (BMBF) [24]. The AEQUIPA research network is comprised of six subprojects and conducts theory-based and participatory empirical research on different aspects of PA and healthy aging in the northwestern part of Germany [24]. One of the aims of the network is to develop, implement, and evaluate PA interventions for the primary prevention of chronic diseases in adults aged at least 65 years. The first 3-year funding phase was completed in January 2018 [25]. The topic of the research remains the same in the second funding phase (2018-2021); however, in this phase, additional aims are to intensify community participation to reach physically inactive adults, use appropriate technology for PA promotion in the population of older adults, and disseminate and transfer PA interventions as well as results of both funding phases. This study (PROMOTE II) is one of the 6 subprojects of the entire network and builds on knowledge gained during the preceding study (PROMOTE I) [19,25,26].

In the previous study, the research aims were to develop and test 2 individually tailored IT-based interventions for the promotion of a physically active lifestyle in adults aged 65-75 years in a community-based intervention trial [25]. Intervention effects on physical, psychological, and cognitive indicators for healthy aging were examined by comparing 2 intervention groups to a delayed-intervention control group [26]. Results of this trial are reported elsewhere [19,26]. Briefly, the proportion of already active intervention participants at baseline was relatively high. At follow-up, they reported increased social-cognitive predictors for behavior change, but no significant increases in the primary outcome of MVPA when compared with the control group. Also, the dropout rate was higher in the group invited to use PA trackers in addition to the IT-based intervention compared with the IT-only intervention suggesting that individuals with little technological experience might have been randomized to an intervention that they found too difficult to use. In addition, participants in both intervention groups requested printed diaries to track their PA during times that they had limited access to computers. Thus, reactions of the target population during the first funding phase suggested that a proportion of participants prefers a print-based intervention or would appreciate an app for accessing a PA diary on their smartphone or tablet.

Study Aims

Hence, based on previous findings of the literature and knowledge gained during the preceding trial, this study aimed to adapt and simplify the IT-based intervention of the first funding phase further to improve usability and develop a simple print-based intervention that initially inactive participants with little affinity to technology find easy to interact with; investigate the implementation, feasibility, and use of 2 interventions (IT-based vs print-based) as well as changes in PA among older adults (aged ≥60 years) in a randomized trial with a crossover design over the course of 9 months; examine the role of personal

preferences for different delivery modes with regard to intervention effectiveness; and explore associations between changes in PA and possible changes in physical fitness and cognitive capacity in a pooled sample of participants from both phases of funding.

Selection of Communities for the Study, Participants, and Procedures

We selected 14 community districts in 2 geographically different regions (northwest and northeast) of the city of Bremen, Germany: Burgdamm, Lesum, St. Magnus, Vegesack, Schönebeck, Aumund-Hammersbeck, Rönnebeck, Radio Bremen, Riensberg, Gartenstadt Vahr, Neue Vahr Südwest, Oberneuland, Ellener Feld, and Blockdiek. They were chosen because there are study centers in these districts that can be easily reached by the target group and the project team had already established prior liaisons with stakeholders in these areas of the city who could facilitate community involvement in the implementation of the intervention.

Names and addresses of men and women aged ≥60 years residing in the chosen community districts were drawn from the records of the residents' registration office. Subsequently, they were invited to participate in the study via mail. Reminders were sent out in case of no response after 2 weeks. The study was also publicized in local newspaper articles and mentioned during talks that targeted older adults by researchers in the team. Individuals made aware of the study through this channel could contact the research team directly to be screened for eligibility. Eligibility for study participation was determined during computer-assisted telephone interviews with trained study nurses based on the inclusion and exclusion criteria.

Inclusion and Exclusion Criteria

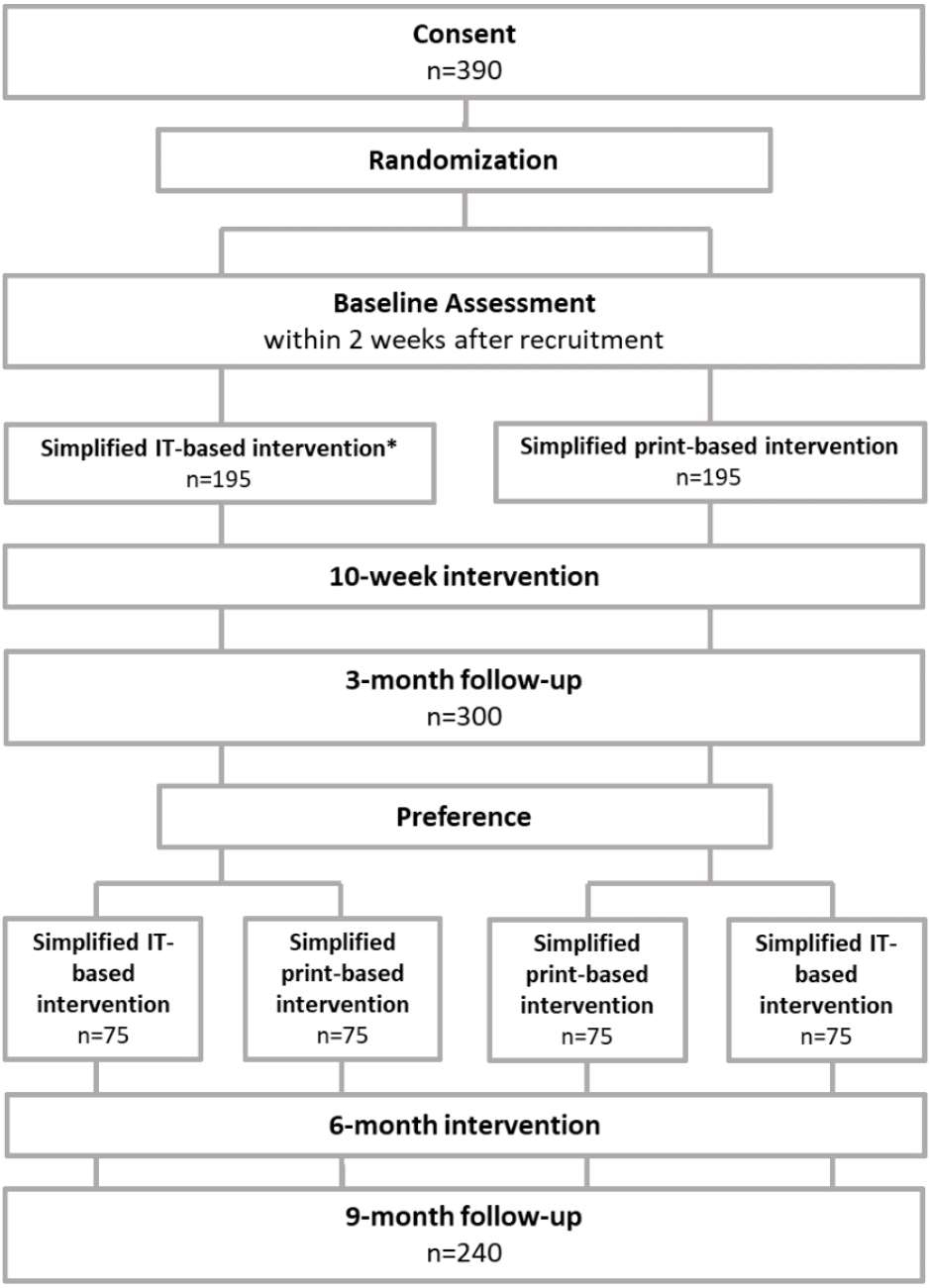
Men and women were eligible for study participation if they were aged ≥60 years (there was no upper age limit), lived independently (ie, in own apartment or room without assisted living or regular home nursing), and provided informed consent to participate in the study. Additional inclusion criteria were basic knowledge of German, the ability to walk without a walking aid and participate in study assessments and weekly group meetings without external assistance, and no planned long absence (ie, for more than 2 weeks). Another precondition was the availability of a device with internet access in the household, regular access to the device, and the ability to use it.

Individuals were excluded from the study if they reported that they had already been regularly physically active for more than 1 year and had reached the intervention target (ie, the recommended 150 minutes of moderate-intensity endurance training per week). They were also excluded if they reported having participated in the preceding study (PROMOTE I), had planned a vacation lasting longer than 2 weeks during the intervention period, had been medically prohibited to be physically active, displayed severe visual or cognitive impairment (Mini-Mental-State Examination Score ≤14) or other impairments (eg, due to a stroke, transient ischemic attack, brain surgery, or neurological diseases, such as Alzheimer's disease, multiple sclerosis, or Parkinson's disease), had an implanted device (eg, pacemaker, implanted hearing aid), or experienced occasional syncope. Additional exclusion criteria were a lack of medical clearance because of fractures or surgeries in the past 6 months that could potentially constrain study participation; report of severe diseases of the cardiovascular system (eg, cardiac arrhythmia, heart failure, or nonmedicated hypertension) or respiratory system (eg, COPD, severe asthma); or severe limitations due to arthritis or osteoarthritis in the legs or severe osteoporosis or acute injuries of the spine. Individuals with diabetes diagnosed less than 6 months prior and without medical clearance were also excluded from the study.

Study Design

After determination of eligibility, study participants were randomized to either an IT-based intervention or a print-based intervention (see Figure 1). In the first group, 30% were randomly selected to receive a PA tracker in addition to access to the website. Weekly time slots were randomly assigned to both intervention groups. Then, we had participants choose appointments according to their personal time constraints without knowing which intervention group they were assigned to. All intervention groups were offered weekly group sessions, and participants were encouraged to attend each of them. Two weeks before the intervention started, participants in all groups attended an introductory event where they were informed about the study background and procedures. On this day, data collection also started, and participants received the activity tracker and questionnaires for the first time (T0). They were also briefed about their randomization to one of the intervention groups and the existence of the other groups. Therefore, they were not blinded. Participants were given the choice to remain in their intervention group or to cross over to the other group after 3 months (and a 10-week intervention period). With that, a crossover design with the following possible combinations was generated: IT-IT, print-print (matched), or IT-print and print-IT. Preferences regarding the intervention material assessed retrospectively at the 9-month follow-up will be taken into account when analyzing intervention effects later in the study. Research staff members conducting the study were not blinded, and the statistician will not be blinded when analyzing the data.

Figure 1. Study design.



Analytic Strategy

Effects of time and group (main outcomes: subjective and objective measures of PA) will be examined in multivariate analyses. In addition, covariates, such as gender, age, and preferences regarding intervention material, and individual contextual factors (eg, living or social environment) will be taken into account. Furthermore, we will analyze how the intervention enables study participants to change their behavior by running mediation analyses. The estimated target sample size of 390 participants for this study is based on the results of 2 recently published studies on PA in adults [27,28] that reported a mean 111 weekly minutes (SD 116.8 weekly minutes) of MVPA and an intervention effect of an increase in MVPA of 77 minutes per week. However, in a conservative (or realistic) approach, we assumed that the mean increase in MVPA in our study will reach only 40 minutes/week. To detect such a difference between the intervention and control groups with a power of 80% at an alpha of .05 (two-sided), a net sample size of 150 subjects for each group is required. Assuming a loss to follow-up from T0 to T1 of 20-25% (ie, ca. 23%), it is necessary to initially recruit 390 study participants (ie, 195 per group at baseline). Assuming another 20% loss to follow-up from T1 to T2, we expect that approximately 240 individuals at T2 will self-select to one of the 4 groups: IT-IT, print-print, IT-print, or print-IT. Should participants self-select equally into each group (n=60), we will still be able to detect the mean difference of 77 minutes/week [28] between each group, by means of an appropriate post hoc test that is adjusted for multiple comparisons.

Measures

The primary outcome is change in PA, which is assessed at baseline, 3 months, and 9 months using triaxial accelerometers worn at the right hip during the day over the course of 7 days and a self-administered questionnaire. Participants received the self-administered questionnaire and accelerometer at their introductory event at baseline (T0), after 3 months (T1) at their tenth group meeting, and after 9 months (T2) at their last monthly meeting. Further, due to cost issues, a random subsample of 114 participants (equally distributed across the intervention groups) underwent additional anthropometric, physical fitness, and cognitive tests to assess secondary outcomes at all 3 time points. Physical fitness was assessed using gait speed via a 4-meter walk test [29], handgrip strength using a dynamometer [30], and cardiovascular fitness using the 2-minute step test [31]. Additionally, participants' weight was assessed using a scale. Attention and inhibition were assessed using a cognitive test (Simon task) [32].

Data from the objective measures of physical fitness and cognitive performance (Mini Mental State Examination and executive function test) will be preprocessed using customized MATLAB routines and analyzed using multivariate generalized linear models. Data for physical fitness and cognitive performance assessed at baseline and after the 10-week intervention during the first funding phase and in the subsample of participants during the second funding phase will be pooled to increase power for data analysis and to meet sample heterogeneity. Further, the larger sample size will enable us to conduct subgroup analyses with respect to motivational stage, activity behavior, and social engagement.

All participants underwent a short version of the Mini Mental State Examination during their first weekly group meeting [33]. Motivational stages to engage in recommended PA (endurance, strength, flexibility, and balance training), social-cognitive predictors for behavior change, and psychosocial factors (eg, quality of life, depression) were assessed in the self-administered questionnaire (for further detail on the instruments included in the questionnaire, see Table 1).

Table 1. Measures assessed in the self-administered study questionnaire.

Outcome measure	Instrument/scale	Time of assessment
Physical activity		
Physical activity	Godin Leisure-Time Exercise Questionnaire (modified) [34]	T0[a], T1[b], T2[c]
Recommended endurance, strength, balance, and flexibility training at follow-up	Self-generated items	T1, T2
Stage of change regarding physical activity (endurance, strength, and balance + flexibility training, respectively)	HAPA[d], stage of change, 3 items [35,36]	T0, T1, T2
Intention to engage in physical activity	HAPA, intention, 2 items [36-38]	T0, T1, T2
Self-efficacy regarding physical activity	HAPA, self-efficacy, 5 items [36,38]	T0, T1, T2
Positive and negative outcome expectancies regarding physical activity	HAPA, outcome expectancies, 4 items [36,38,39]	T0, T1, T2
Planning for physical activity	HAPA, planning, 6 items [38,40]	T0, T1, T2
Habit strength regarding physical activity	Self-Report Habit Index, 2 items [41]	T0, T1, T2
Physical self-description	PSDQ[e] [42]	T0, T1, T2
Physical environment		
Physical activity and neighborhood environment	Physical Activity Neighborhood Environment Scale [43]	T0
Walking environment	Neighborhood Scales, walking environment [44]	T0
Social support, social activities		
Social support for engaging in physical activity	Activity-related support by family and friends (modified) [45,46]	T0, T1, T2
Social activities	Florida Cognitive Activities Scale (modified) [47,48]	T0, T1, T2
Health behaviors		
Subjective age	Difference score (perceived physical age - chronological age = subjective physical age) [49]	T0, T1, T2
Subjective health status	WHOQOL-BREF[f], 1 item [50,51]; SF-36[g], 1 item [52]	T0, T1, T2
Health-related quality of life	EQ-5D-3L[h] [53,54]	T0, T1, T2
Objective health	Diseases and medication use (modified) [55]	T0
Falls	EFST[i] (modified) [56]	T0, T1, T2
Fear of falling	GFFM[j] [57]	T0, T1, T2
Diet	FFQ[k] (modified) [58]	T0, T1, T2
Alcohol consumption	AUDIT-C[l] [59]	T0, T1, T2
Smoking behavior	Smoking Behavior Questionnaire, 1 item [60]	T0, T1, T2
Stage assessment of smoking behavior	HAPA, stage of change [35,36]	T0, T1, T2
Quality of life and well-being		
Quality of life	WHOQOL-BREF, 1 item [50,51]	T0, T1, T2
Depression	CES-D[m] [61]	T0, T1, T2
Previous experiences with technology		
Use of computers/smartphones/apps	Self-generated items	T0, T1, T2
Technology commitment	Technology Commitment Scale [62]	T0, T1, T2
Use of, acceptance of, and satisfaction with interventions		
Use and acceptance of various components of the interventions (website and printed material), attendance of the offered group sessions, and overall satisfaction with the interventions	Self-generated items	T1, T2

Outcome measure	Instrument/scale	Time of assessment
Preference regarding intervention material at baseline (retrospective)	Self-generated item	T2
Reasons for crossing over or not crossing over after 3 months	Self-generated items	T2

[a]Baseline.
[b]3-month follow-up.
[c]9-month follow-up.
[d]HAPA: Health Action Process Approach.
[e]PSDQ: Physical Self-Description Questionnaire.
[f]WHOQOL-BREF: World Health Organization Quality of Life-BREF.
[g]SF-36: Short Form 36.
[h]EQ-5D-3L: 3-level version of the EQ-5D.
[i]EFST: Elderly Fall Screening Test.
[j]GFFM: Geriatric Fear of Falling Measurement.
[k]FFQ: Food Frequency Questionnaire.
[l]AUDIT-C: Alcohol Use Disorders Identification Test Short Version.
[m]CES-D: Centers for Epidemiologic Studies Depression Scale.

Anthropometric information, as self-reported in the questionnaire, included height in cm (only T0) and weight in kg (all time points). Sociodemographic information was assessed at baseline only, using items of the German Health Interview and Examination Survey for Adults [63] for date and country of birth, gender, native language, family and relationship status, living situation, education, employment status, and monthly net household income. Employment was assessed using one item of a questionnaire for assessing seniors' demographic and sociostructural data in Germany [64]. In addition, the self-administered questionnaire at T1 and T2 includes self-generated items regarding adherence to the PA recommendations (weekly minutes of endurance, balance, and flexibility training and weekly units of muscle strength training per muscle group) and attendance of the offered group sessions. Furthermore, use and satisfaction with the intervention material and content are assessed with the following self-generated items: frequency of general use; use of different intervention components reporting using a 5-point Likert scale from "never" to "daily"; perceived helpfulness of intervention components, intervention content, and structure of weekly meetings rated on a 5-point Likert scale ranging from "not helpful at all" to "very helpful"; potential knowledge gained and perceived benefits of PA using "yes" and "no" response options; and whether participants would recommend the intervention to family and friends, answered using open-ended questions. Data on the reasons for dropping out of the study and the reasons for crossing over or not crossing over to the respective other intervention group were also collected.

Interventions

To simplify and further develop the IT-based intervention implemented during the first funding phase and to translate it into a print-based intervention, 9 focus groups consisting of former participants of the previous intervention (n=1), other members of the target group (n=32), and stakeholders in communities such as members of senior citizen organizations and advisory boards (n=12) were held in May and June 2018. Participants of these focus groups were recruited via a press release and already established contacts to the stakeholders. Focus group discussions consisted of two parts. First, intervention materials used in the previous study and materials and health brochures for PA promotion developed by other researchers; federal agencies, such as the Federal Ministry of Health (BMG); or health insurance agencies were discussed with regard to their suitability for PA promotion in inactive older adults. In this context, the participants were invited to vote for the material they liked most and to justify their decision. Second, using the World Café method, participants were asked to discuss the following questions in small groups of 4-5 participants:

1. From your perspective, what are the key aspects of healthy aging?
2. Which topics related to healthy aging should be addressed in an intervention in order to reach physically inactive older adults?
3. What do inactive older adults need to be able to successfully participate in a PA intervention?
4. Which barriers need to be overcome and what would be helpful to these participants?
5. What would be motivating for inactive older adults for maintaining a healthy lifestyle?

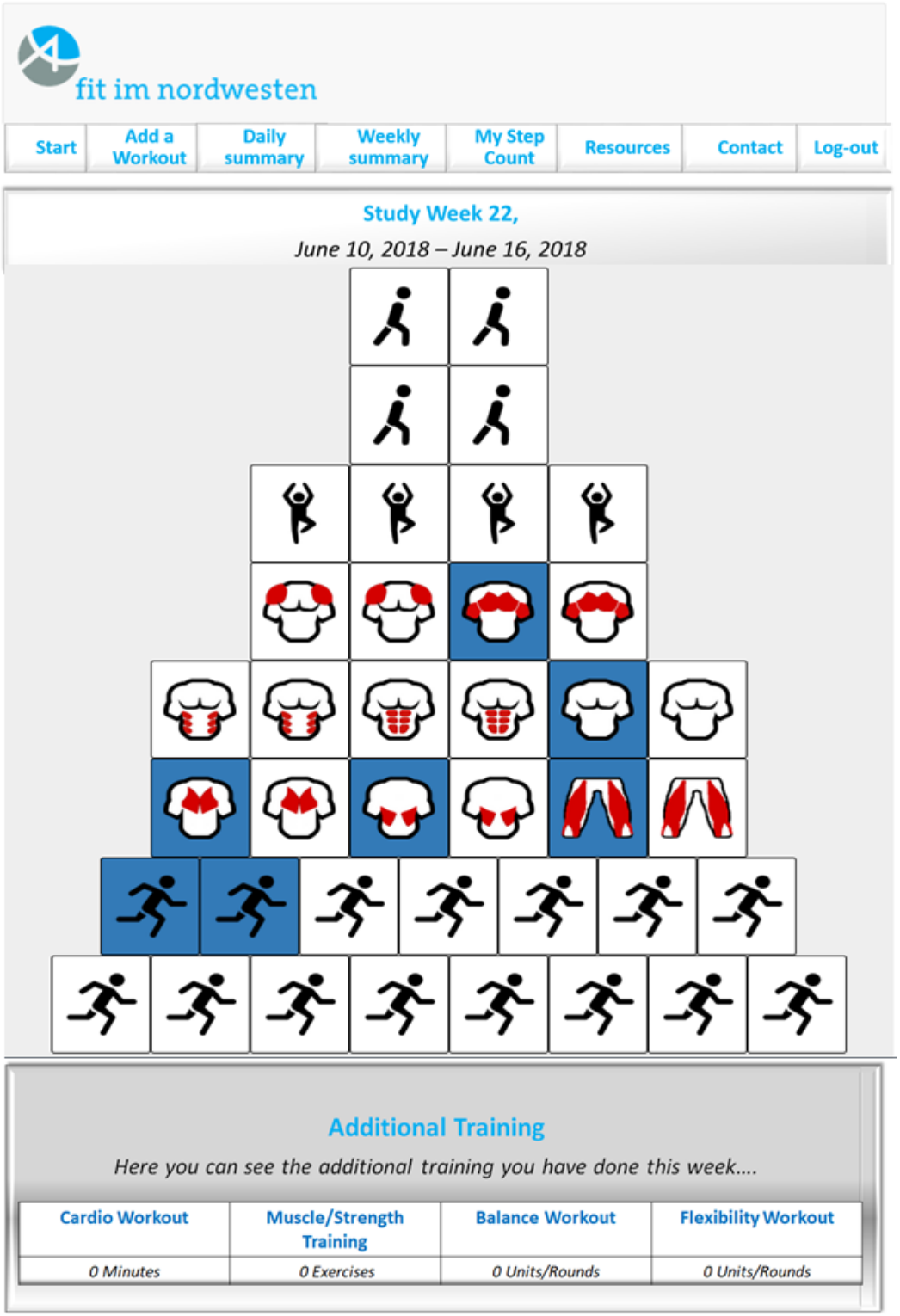
The focus group discussions and votes for or against certain intervention materials were protocolled by the moderating researchers, and the resulting recommendations and requirements for the intervention material were applied when refining the existing IT-based intervention and when translating the IT-based intervention into a print-based intervention. Results consisted of recommendations regarding font size and color, amount of text to read, appropriateness of content (eg, inclusion of recommendations regarding diet, sleep hygiene, or pain management), and considerations with regard to the target group (eg, mobility, physical limitations, and the role of social support, loneliness, and wellbeing).

Both interventions are based on self-regulation theory [65,66] and principles of behavior change (eg, shaping knowledge,

feedback and monitoring, goals and planning, social support, comparison of behavior, rewards) [67] and are designed to facilitate a physically active lifestyle by promoting regular self-monitoring of PA. Participants of both the IT-based and print-based interventions received PA recommendations based on WHO recommendations for this age group regarding endurance training (at least 150 minutes per week of moderate exercise, at least 75 minutes of intensive exercise, or a combination of both, in bouts of at least 10 minutes), strength training (at least 2 units per week for the 8 main muscle groups on non-consecutive days), and balance and flexibility training (at least 4 units per week for 5 minutes) [2]. Depending on gender, participants were provided with different brochures (online and offline) outlining exercises for different difficulty levels and displaying pictures of male or female older adults modeling these exercises.

They also received a weekly PA diary in the form of a pyramid to subjectively track their PA behavior. Intervention material was printed for the print-based intervention group, while participants in the IT-based group received access to the exercise brochure and diary on a website and were provided the opportunity to download a smartphone app containing similar features. On the website or Android app, as well as in the printed diary, weekly feedback on whether participants reached PA goals was provided by displaying the amount of minutes or units exercised in the according week, as well as units required to reach the goal (ie, WHO recommendations for moderate exercise time, flexibility, and strength training). In addition to the weekly feedback, the website and Android app also provided participants a daily overview of PA. In the daily overview, participants had access to more detailed information about the exercises for the individual muscle groups, whereas in the weekly overview (in the IT-based and print-based interventions), only information about the progress towards the weekly goals was provided (see Figure 2). The app also allowed participants to record exercises without an active internet connection and synchronized recordings with the website, once the device was connected to the internet again. Participants in the IT-based intervention group who used the PA tracker in addition to the website or app also monitored their PA behavior objectively via synchronization of data regarding daily steps with the website. Data from the PA trackers will be analyzed based on methods presented in Meyer et al [68] to gain insights into participants' PA behavior in everyday life during the time between assessments.

Figure 2. Weekly overview of physical activity performed, as provided to participants in the IT-based intervention group.



Paralleling the 10-week IT-based and print-based interventions, weekly group sessions with up to 25 participants per group were offered separately to both intervention groups. Trained student assistants led these sessions in which participants performed endurance, strength, balance, and flexibility exercises in groups or went for walks in the communities. Each session was 90 minutes long and designed to include 60 minutes of exercise, 10 minutes of answering questions, and 20 minutes of discussion of topics surrounding healthy aging that changed every week (eg, social support, relaxation, pain management). During their first weekly group meeting, participants received the necessary equipment (printed material or information to access the website and/or a Fitbit) and a comprehensive introduction on how to use the equipment and materials. After the first follow-up, interventions continued for another 6 months. After the initial 10 weeks of the interventions, the weekly group meetings were replaced with monthly events held over the course of the following 6 months. During these events, participants of all

intervention groups were offered workshops and lectures on lifestyle-related topics, such as healthy nutrition in older age, overcoming loneliness, and strategies for developing healthy habits.

We will develop a toolbox during this project that will include both IT-based and print-based interventions and all related intervention materials that can be used later for long-term implementation and dissemination via different stakeholder groups (for further detail, please see the Discussion).

Ethics Statement and Consent

Ethical approval was obtained from the Medical Association in Bremen (RA/RE-635, on July 3, 2018). The study was registered with the German Clinical Trials Register on January 10, 2019 – number DRKS00016073. All study participants were fully informed about the study and were requested to give informed consent.

Results

In this study, we refined the previously developed IT-based intervention and translated it into a simple print-based intervention during the development phase from February 2018 to December 2018. The results of the qualitative focus groups held in May and June 2018 were analyzed and used to refine the existing intervention and to develop the print-based intervention. During the implementation phase of the study (baseline: January to April 2019; 3-month follow-up: April to July 2019; 9-month follow-up: September 2019 to January 2020), we expect to observe significant increases in PA at the 3-month follow-up after participation in both the print-based and IT-based interventions, in comparison with baseline, and a larger increase in PA in the 2 intervention groups than in the delayed-intervention control group of the preceding study (PROMOTE I).

We further expect that by providing some degree of choice to study participants 3 months into this study, we will be able to reduce loss to follow-up and improve long-term program adherence, when compared with the previous study. Further, we assume that participants with a ≥25% increase in PA level from baseline to the end of the follow-up period in the pooled sample across the 2 funding phases (ie, PROMOTE I sample and the subsample in PROMOTE II) will display significant improvements in indicators for healthy aging, such as cognitive function. Data collection was completed on January 31, 2020. Data cleaning and analysis started in February 2020. We expect to publish the first results of the study by the end of the funding period (January 2021).

Discussion

This study will provide answers regarding the acceptance and effectiveness of IT-based vs print-based interventions for promoting uptake and maintenance of regular PA in initially inactive individuals aged 60 years and older. Further, we hope to generate the first results regarding the role of individual preferences for various intervention delivery modes in this study and the potential of a preference-based crossover design. In addition, this study will provide insights into the needs and demands of vulnerable groups (ie, inactive older adults). We will be able to identify “user groups” with regard to an affinity for specific intervention components.

This topic is of particular relevance in Germany, as 29% of the population is projected to be older than 65 years by 2030 [69]. Regular PA is a key contributor to healthy aging. Results regarding the effects of the interventions on PA and other health outcomes, such as quality of life, in initially inactive older adults over a relatively long period and their preferences for different intervention modes will be valuable to various stakeholder groups. For example, the work of stakeholders actively involved in community-based networks and senior citizen associations promoting population-based strategies for active aging will be informed by our results and implementation experiences. Further, the Prevention Law, which was passed in 2016 in Germany [70], mandated that health insurance agencies invest in health promotion and primary prevention in various contexts, including communities. Thus, our results will improve the existing evidence base on the effectiveness and implementation of community-based interventions in Germany and will support efforts to anchor evidence-based PA interventions in community structures and organizations via an allocation of permanent health insurance funds. To facilitate the rollout of the interventions, a toolbox for these stakeholder groups will be developed that includes PA assessment and monitoring tools allowing individuals to track their PA and helping health care professionals support their clients’ behavior change. We will also provide strategies to facilitate behavior adoption and maintenance and instructions to use personalized digital tools as standalone interventions without face-to-face assistance.

Last, complementing existing health and preventive care with eHealth intervention approaches, such as mobile apps for PA promotion, is in line with the current eHealth initiative of the Federal Ministry of Health [71]. The project is embedded in multidisciplinary research aimed at better understanding how changes in individual behavior and the environment and technology can help promote healthy aging. Our findings will inform the future development of complex interventions combining both local and regional policy changes aimed at promoting environmental changes and technology-based interventions targeting individual behavior change. Our experiences regarding this population’s data protection concerns will be communicated to policy makers currently involved in the eHealth initiative and will support the identification of suitable approaches to deal with these concerns.

Despite the advantages of the study design and objective PA measurement, several limitations can be identified. First, this study design does not include a control group, which might limit the interpretations that can be drawn from the potentially positive intervention effects. Second, the target population consists of inactive older adults, which are difficult to recruit. Third, the retrospective assessment of the preference for a certain intervention delivery mode, which could result in recall bias. However, we chose to retrospectively assess preferences because we anticipated disappointment (and possibly dropouts) if a person was not randomized to their preferred intervention group at baseline.

To conclude, this study will provide insights into the acceptance and effectiveness of an IT-based vs a print-based intervention for the promotion of PA in initially inactive individuals aged ≥60 years. In addition to answering the main research questions, we expect to obtain a better understanding of the interactions between numerous contextual factors and PA in this population. Results from this study will inform the work of various stakeholder groups actively involved in PA promotion at the population level and in different contexts and will support the shaping of new policies regulating the future design and implementation of preventive eHealth interventions for active aging.

Acknowledgments

This study is funded by the German Federal Ministry of Education and Research (BMBF; project numbers 01EL1822A, 01EL1822F, 01EL1822l, 01EL1822C). The content of this article only reflects the authors' views, and the funder is not liable for any use that may be made of the information contained therein.

Data Availability

After completion of the study, data can be shared with other researchers upon request but only in research collaborations with researchers of the study.

Authors' Contributions

CRP, MP, TR, and SL drafted the manuscript. CRP, MP, TR, SL, HP, CVR, JM, and KvH made substantial contributions to the conception and design of the study. All authors read, critically revised, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer Review Report - AEQUIIPA.
[PDF File (Adobe PDF File), 98 KB - resprot_v9i4e15168_app1.pdf]

References

1. World Health Organization. Noncommunicable diseases country profiles. Geneva: World Health Organization; 2018.
2. World Health Organization. Global recommendations on physical activity for health. Geneva: WHO Press; 2010:978924159997.
3. Hong S, Hughes S, Prohaska T. Factors affecting exercise attendance and completion in sedentary older adults: a meta-analytic approach. J Phys Act Health 2008 May;5(3):385-397. [doi: 10.1123/jpah.5.3.385] [Medline: 18579917]
4. Byberg L, Melhus H, Gedeborg R, Sundström J, Ahlbom A, Zethelius B, et al. Total mortality after changes in leisure time physical activity in 50 year old men: 35 year follow-up of population based cohort. Br J Sports Med 2009 Jul;43(7):482. [Medline: 19581403]
5. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, et al. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1435-1445. [doi: 10.1249/mss.0b013e3180616aa2] [Medline: 17762378]
6. Finger J, Mensink G, Lange C, Manz K. Gesundheitsfördernde körperliche Aktivität in der Freizeit bei Erwachsenen in Deutschland. Journal of Health Monitoring 2017;2(2):37-44. [doi: 10.17886/RKI-GBE-2017-027]
7. Rütten A, Pfeifer K. Nationale Empfehlungen für Bewegung und Bewegungsförderung. Erlangen-Nürnberg: FAU; 2016:129.
8. Krug S, Jordan S, Mensink GBM, Müters S, Finger J, Lampert T. [Physical activity: results of the German Health Interview and Examination Survey for Adults (DEGS1)]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013 May;56(5-6):765-771. [doi: 10.1007/s00103-012-1661-6] [Medline: 23703496]
9. Hagger MS, Weed M. DEBATE: Do interventions based on behavioral theory work in the real world? Int J Behav Nutr Phys Act 2019 Dec 25;16(1):36 [FREE Full text] [doi: 10.1186/s12966-019-0795-4] [Medline: 31023328]
10. Prestwich A, Sniehotta FF, Whittington C, Dombrowski SU, Rogers L, Michie S. Does theory influence the effectiveness of health behavior interventions? Meta-analysis. Health Psychol 2014 May;33(5):465-474. [doi: 10.1037/a0032853] [Medline: 23730717]
11. Amireault S, Baier JM, Spencer JR. Physical Activity Preferences Among Older Adults: A Systematic Review. J Aging Phys Act 2018 Oct 25;27(1):1-12. [doi: 10.1123/japa.2017-0234] [Medline: 29283793]
12. Richards J, Hillsdon M, Thorogood M, Foster C. Face-to-face interventions for promoting physical activity. Cochrane Database Syst Rev 2013 Sep 30;9(9):CD010392. [doi: 10.1002/14651858.CD010392.pub2] [Medline: 24085592]
13. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]

14. Short CE, James EL, Plotnikoff RC, Girgis A. Efficacy of tailored-print interventions to promote physical activity: a systematic review of randomised trials. Int J Behav Nutr Phys Act 2011 Oct 17;8:113 [FREE Full text] [doi: 10.1186/1479-5868-8-113] [Medline: 21999329]
15. Joseph RP, Durant NH, Benitez TJ, Pekmezi DW. Internet-Based Physical Activity Interventions. Am J Lifestyle Med 2014 Dec;8(1):42-68 [FREE Full text] [doi: 10.1177/1559827613498059] [Medline: 25045343]
16. Jahangiry L, Farhangi MA, Shab-Bidar S, Rezaei F, Pashaei T. Web-based physical activity interventions: a systematic review and meta-analysis of randomized controlled trials. Public Health 2017 Nov;152:36-46. [doi: 10.1016/j.puhe.2017.06.005] [Medline: 28734170]
17. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: A systematic review. Prev Med 2018 Mar;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]
18. Vandelanotte C, Duncan MJ, Maher CA, Schoeppe S, Rebar AL, Power DA, et al. The Effectiveness of a Web-Based Computer-Tailored Physical Activity Intervention Using Fitbit Activity Trackers: Randomized Trial. J Med Internet Res 2018 Dec 18;20(12):e11321 [FREE Full text] [doi: 10.2196/11321] [Medline: 30563808]
19. Muellmann S, Buck C, Voelcker-Rehage C, Bragina I, Lippke S, Meyer J, et al. Effects of two web-based interventions promoting physical activity among older adults compared to a delayed intervention control group in Northwestern Germany: Results of the PROMOTE community-based intervention trial. Prev Med Rep 2019 Sep;15:100958 [FREE Full text] [doi: 10.1016/j.pmedr.2019.100958] [Medline: 31410347]
20. Torgerson DJ, Klaber-Moffett J, Russell IT. Patient preferences in randomised trials: threat or opportunity? J Health Serv Res Policy 1996 Oct;1(4):194-197. [Medline: 10180870]
21. Preference Collaborative Review Group. Patients' preferences within randomised trials: systematic review and patient level meta-analysis. BMJ 2008;337:a1864 [FREE Full text] [Medline: 18977792]
22. Short CE, Vandelanotte C, Duncan MJ. Individual characteristics associated with physical activity intervention delivery mode preferences among adults. Int J Behav Nutr Phys Act 2014;11(1):25 [FREE Full text] [doi: 10.1186/1479-5868-11-25] [Medline: 24568611]
23. Alley S, Schoeppe S, Guertler D, Jennings C, Duncan MJ, Vandelanotte C. Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. BMJ Open 2016 Dec 07;6(7):e011243 [FREE Full text] [doi: 10.1136/bmjopen-2016-011243] [Medline: 27388359]
24. Forberger S, Bammann K, Bauer J, Boll S, Bolte G, Brand T, et al. How to Tackle Key Challenges in the Promotion of Physical Activity among Older Adults (65+): The AEQUIPA Network Approach. Int J Environ Res Public Health 2017 Dec 04;14(4) [FREE Full text] [doi: 10.3390/ijerph14040379] [Medline: 28375177]
25. Muellmann S, Bragina I, Voelcker-Rehage C, Rost E, Lippke S, Meyer J, et al. Development and evaluation of two web-based interventions for the promotion of physical activity in older adults: study protocol for a community-based controlled intervention trial. BMC Public Health 2017 Dec 25;17(1):512 [FREE Full text] [doi: 10.1186/s12889-017-4446-x] [Medline: 28545506]
26. Ratz T, Lippke S, Muellmann S, Peters M, Pischke CR, Meyer J, et al. Effects of Two Web-Based Interventions and Mediating Mechanisms on Stage of Change Regarding Physical Activity in Older Adults. Appl Psychol Health Well Being 2020 Mar;12(1):77-100. [doi: 10.1111/aphw.12174] [Medline: 31332957]
27. Van Holle V, De Bourdeaudhuij I, Deforche B, Van Cauwenberg J, Van Dyck D. Assessment of physical activity in older Belgian adults: validity and reliability of an adapted interview version of the long International Physical Activity Questionnaire (IPAQ-L). BMC Public Health 2015 Apr 28;15:433 [FREE Full text] [doi: 10.1186/s12889-015-1785-3] [Medline: 25928561]
28. Wijsman CA, Westendorp RG, Verhagen EA, Catt M, Slagboom PE, de Craen AJ, et al. Effects of a web-based intervention on physical activity and metabolism in older adults: randomized controlled trial. J Med Internet Res 2013;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
29. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [Medline: 8126356]
30. Bohannon RW, Schaubert KL. Test-retest reliability of grip-strength measures obtained over a 12-week interval from community-dwelling elders. J Hand Ther 2005;18(4):426-7, quiz 428. [doi: 10.1197/j.jht.2005.07.003] [Medline: 16271690]
31. Rikli R, Jones C. Functional fitness normative scores for community-residing older adults, ages 60-94. J Aging Phys Act 1999;7(2):162-181. [doi: 10.1123/japa.7.2.162]
32. Bialystok E, Craik F, Luk G. Cognitive control and lexical access in younger and older bilinguals. J Exp Psychol Learn Mem Cogn 2008 Jul;34(4):859-873. [doi: 10.1037/0278-7393.34.4.859] [Medline: 18605874]
33. Folstein M, Folstein S, White T, Messer M. MMSE®-2 Mini-Mental® State Examination, 2nd Edition. Lutz, Florida: PAR; 2010.
34. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]

35. Lippke S, Fleig L, Pomp S, Schwarzer R. Validity of a stage algorithm for physical activity in participants recruited from orthopedic and cardiac rehabilitation clinics. Rehabil Psychol 2010 Nov;55(4):398-408. [doi: 10.1037/a0021563] [Medline: 21171799]
36. Lippke S, Ziegelmann JP, Schwarzer R, Velicer WF. Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. Health Psychol 2009 Mar;28(2):183-193 [FREE Full text] [doi: 10.1037/a0012983] [Medline: 19290710]
37. Nigg CR. There is more to stages of exercise than just exercise. Exerc Sport Sci Rev 2005 Jan;33(1):32-35. [Medline: 15640718]
38. Schwarzer R, Schuz B, Ziegelmann JP, Lippke S, Luszczynska A, Scholz U. Adoption and maintenance of four health behaviors: theory-guided longitudinal studies on dental flossing, seat belt use, dietary behavior, and physical activity. Ann Behav Med 2007 Apr;33(2):156-166. [doi: 10.1007/bf02879897] [Medline: 17447868]
39. Lippke S, Ziegelmann JP, Schwarzer R. Behavioral Intentions and Action Plans Promote Physical Exercise: A Longitudinal Study with Orthopedic Rehabilitation Patients. Journal of Sport and Exercise Psychology 2004 Sep;26(3):470-483. [doi: 10.1123/jsep.26.3.470]
40. Sniehotta FF, Schwarzer R, Scholz U, Schüz B. Action planning and coping planning for long-term lifestyle change: theory and assessment. Eur. J. Soc. Psychol 2005 Jul;35(4):565-576. [doi: 10.1002/ejsp.258] [Medline: 25855820]
41. Verplanken B, Orbell S. Reflections on Past Behavior: A Self-Report Index of Habit Strength1. J Appl Social Pyschol 2003 Jun;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
42. Stiller J, Würth S, Alfermann D. Die Messung des physischen Selbstkonzepts (PSK). Zeitschrift für Differentielle und Diagnostische Psychologie 2004 Jan;25(4):239-257. [doi: 10.1024/0170-1789.25.4.239]
43. Sallis JF, Kerr J, Carlson JA, Norman GJ, Saelens BE, Durant N, et al. Evaluating a brief self-report measure of neighborhood environments for physical activity research and surveillance: Physical Activity Neighborhood Environment Scale (PANES). J Phys Act Health 2010 Jul;7(4):533-540. [doi: 10.1123/jpah.7.4.533] [Medline: 20683096]
44. Mujahid MS, Diez Roux AV, Morenoff JD, Raghunathan T. Assessing the measurement properties of neighborhood scales: from psychometrics to ecometrics. Am J Epidemiol 2007 Apr 15;165(8):858-867. [doi: 10.1093/aje/kwm040] [Medline: 17329713]
45. Jackson J, Lippke S, Gray CD. Stage-Specific Prediction of Physical Activity in Orthopaedic Patients after Rehabilitation Treatment. Int J Sport Psychol Dec 2011;42(6):586-609.
46. Fuchs R. Psychologie und Körperliche Bewegung. Göttingen: Hogrefe; 1997:9783801708764.
47. Schinka JA, McBride A, Vanderploeg RD, Tennyson K, Borenstein AR, Mortimer JA. Florida Cognitive Activities Scale: initial development and validation. J Int Neuropsychol Soc 2005 Jan;11(1):108-116. [doi: 10.1017/S1355617705050125] [Medline: 15686613]
48. Jopp DS, Hertzog C. Assessing adult leisure activities: an extension of a self-report activity questionnaire. Psychol Assess 2010 Mar;22(1):108-120 [FREE Full text] [doi: 10.1037/a0017662] [Medline: 20230157]
49. Wienert J, Kuhlmann T, Fink S, Hambrecht R, Lippke S. Subjective ageplanning strategies for physical activity: brief report about a domain-specific example from a cross-sectional observation study. Int. J. Sport Psychol Oct 2017;48(4):448-458. [doi: 10.7352/IJSP2017.48.448]
50. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [Medline: 15085902]
51. The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med 1998 May;28(3):551-558. [Medline: 9626712]
52. Bullinger M, Kirchberger I. Fragebogen zum Gesundheitszustand: SF-36. Handanweisung. Göttingen: Hogrefe; 1998.
53. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
54. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [Medline: 11491192]
55. Hultsch DF, Hertzog C, Small BJ, Dixon RA. Use it or lose it: engaged lifestyle as a buffer of cognitive decline in aging? Psychol Aging 1999 Jun;14(2):245-263. [doi: 10.1037//0882-7974.14.2.245] [Medline: 10403712]
56. Cwikel JG, Fried AV, Biderman A, Galinsky D. Validation of a fall-risk screening test, the Elderly Fall Screening Test (EFST), for community-dwelling elderly. Disabil Rehabil 1998 May;20(5):161-167. [doi: 10.3109/09638289809166077] [Medline: 9622261]
57. Huang T, Wang W. Comparison of three established measures of fear of falling in community-dwelling older adults: psychometric testing. Int J Nurs Stud 2009 Oct;46(10):1313-1319. [doi: 10.1016/j.ijnurstu.2009.03.010] [Medline: 19394017]
58. Gallois KM, Buck C, Dreas JA, Hassel H, Zeeb H. Evaluation of an intervention using a self-regulatory counselling aid: pre- and post- intervention results of the OPTIMAHL 60plus study. Int J Public Health 2013 Jun;58(3):449-458. [doi: 10.1007/s00038-012-0420-7] [Medline: 23111370]

59. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. Addiction 1993 Jun;88(6):791-804. [Medline: 8329970]
60. Renner B, Schwarzer R. Risk and health behaviors. Documentation of the scales of the research project: Risk Appraisal Consequences in Korea (RACK). Second ed; 2005.
61. Radloff LS. The CES-D Scale: a self-report Depression scale for research in the general population. Applied Psychological Measurement 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
62. Neyer FJ, Felber J, Gebhardt C. Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. Diagnostica 2012 Apr;58(2):87-99. [doi: 10.1026/0012-1924/a000067]
63. Robert Koch Institut. Berlin: Robert-Koch-Institut. 2009. Studie zur Gesundheit Erwachsener in Deutschland (DEGS): Gesundheitsfragebogen ab 65 Jahren URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/degs_projektbeschr.pdf?__blob=publicationFile [accessed 2019-06-25]
64. Berthelsmann Stiftung. Fragebogen TF1: Demografische und sozial-strukturelle Daten. Sozialplanung für Senioren URL: http://www.sozialplanung-senioren.de/frageboegen-bausteine/frageboegen-nach-themenfeldern/index.html?L=0%2F%2F%2F%2F%27tyl [accessed 2019-05-06]
65. Fleig L, Lippke S, Pomp S, Schwarzer R. Intervention effects of exercise self-regulation on physical exercise and eating fruits and vegetables: a longitudinal study in orthopedic and cardiac rehabilitation. Prev Med 2011 Sep;53(3):182-187. [doi: 10.1016/j.ypmed.2011.06.019] [Medline: 21784096]
66. Pomp S, Fleig L, Schwarzer R, Lippke S. Effects of a self-regulation intervention on exercise are moderated by depressive symptoms: A quasi-experimental study. International Journal of Clinical and Health Psychology 2013 Jan;13(1):1-8. [doi: 10.1016/S1697-2600(13)70001-2]
67. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
68. Meyer J, von Holdt K, Beck E, Pischke C, Voelcker-Rehage C. Toy or tool? Activity trackers for the assessment of physical activity in the wild. 2019 Presented at: IEE International Conference on Healthcare Informatics; June 10-13th, 2019; Xian, China. [doi: 10.1109/ichi.2019.8904584]
69. Statista. Prognose des Anteils der Bevölkerung ab 65 Jahren und ab 85 Jahren in Deutschland in den Jahren 2030 und 2060 URL: https://de.statista.com/statistik/daten/studie/196598/umfrage/prognose-des-anteils-der-bevoelkerung-ab-65-jahren-in-deutschland/ [accessed 2020-01-17]
70. Federal Ministry of Health. Promoting preventive healthcare. Berlin, Germany; 2015. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/Publikationen/Ministerium/Broschueren/BMG_Hausbroschuere_2016_EN.pdf [accessed 2020-03-11]
71. Federal Ministry of Health. E-Health Digitalisierung im Gesundheitswesen URL: https://www.bundesgesundheitsministerium.de/e-health-initiative.html [accessed 2020-01-20]

Abbreviations

AEQUIPA: Physical Activity and Health Equity: Primary Prevention for Healthy Ageing
AUDIT-C: Alcohol Use Disorders Identification Test Short Version
CES-D: Centers for Epidemiologic Studies Depression Scale
eHealth: electronic health
EFST: Elderly Fall Screening Test
EQ-5D-3L: 3-level version of the EQ-5D
FFQ: Food Frequency Questionnaire
GFFM: Geriatric Fear of Falling Measurement
HAPA: Health Action Process Approach
IT: information technology
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PSDQ: Physical Self-Description Questionnaire
SF-36: Short Form 36
WHO: World Health Organization
WHOQOL-BREF: World Health Organization Quality of Life-BREF

Edited by G Eysenbach; submitted 25.06.19; peer-reviewed by J Hawkins, SL Tan; comments to author 19.10.19; revised version received 13.12.19; accepted 27.01.20; published 27.04.20.

Please cite as:
Pischke CR, Voelcker-Rehage C, Peters M, Ratz T, Pohlabeln H, Meyer J, von Holdt K, Lippke S
Implementation and Effects of Information Technology-Based and Print-Based Interventions to Promote Physical Activity Among Community-Dwelling Older Adults: Protocol for a Randomized Crossover Trial
JMIR Res Protoc 2020;9(4):e15168
URL: https://www.researchprotocols.org/2020/4/e15168
doi:10.2196/15168
PMID:32338622

Protocol

Efficacy and Safety of Rituximab in Refractory CIDP With or Without IgG4 Autoantibodies (RECIPE): Protocol for a Double-Blind, Randomized, Placebo-Controlled Clinical Trial

Shinobu Shimizu[1*], PhD; Masahiro Iijima[1,2*], MD, PhD; Yuki Fukami[2], MD; Natsuko Tamura[1,3], MSc; Masahiro Nakatochi[1,4], PhD; Masahiko Ando[1], MD, PhD; Ryoji Nishi[2], MD; Haruki Koike[2], MD, PhD; Kenichi Kaida[5], MD, PhD; Michiaki Koga[6], MD, PhD; Takashi Kanda[6], MD, PhD; Hidenori Ogata[7], MD, PhD; Jun-Ichi Kira[7], MD, PhD; Masahiro Mori[8], MD, PhD; Satoshi Kuwabara[8], MD, PhD; Masahisa Katsuno[2], MD, PhD

[1]Department of Advanced Medicine, Nagoya University Hospital, Nagoya, Japan

[2]Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Japan

[3]Center for Integrated Medical Research, Hiroshima University Hospital, Hiroshima, Japan

[4]Department of Nursing, Nagoya University Graduate School of Medicine, Nagoya, Japan

[5]Department of Neurology, Anti-aging and Vascular Medicine, National Defense Medical College, Tokorozawa, Japan

[6]Department of Neurology and Clinical Neuroscience, Yamaguchi University Graduate School of Medicine, Ube, Japan

[7]Department of Neurology, Neurological Institute, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

[8]Department of Neurology, Graduate School of Medicine, Chiba University, Chiba, Japan

[*]these authors contributed equally

Corresponding Author:
Masahiro Iijima, MD, PhD
Department of Neurology
Nagoya University Graduate School of Medicine
65 Tsurumai-cho
Showa-ku
Nagoya, 466-8550
Japan
Phone: 81 52 744 2389
Email: ijama@med.nagoya-u.ac.jp

Abstract

Background: Chronic inflammatory demyelinating polyradiculoneuropathy (CIDP) is an immune-mediated peripheral neuropathy that is currently classified into several clinical subtypes, which are presumed to have different pathogenic mechanisms. Recently, studies identified a subgroup of patients with CIDP who were positive for IgG4 autoantibodies against paranodal proteins, such as neurofascin-155 and contactin-1, who respond poorly to first-line therapies for typical CIDP, including intravenous immunoglobulin therapy.

Objective: This study aims to evaluate the efficacy and safety of intravenous rituximab according to IgG4 autoantibody status in patients with refractory CIDP.

Methods: The Evaluation of the Efficacy and Safety of Rituximab in Refractory CIDP Patients with IgG4 Autoantibodies in the Exploratory Clinical (RECIPE) trial consists of 2 cohorts: a multicenter, placebo-controlled, randomized study cohort of 15 patients with IgG4 autoantibody-positive CIDP (rituximab:placebo = 2:1) and an open-label trial cohort of 10 patients with antibody-negative CIDP. The primary endpoint is improvement in functional outcome assessed using the adjusted Inflammatory Neuropathy Cause and Treatment Disability Scale score at 26, 38, or 52 weeks after the start of treatment with rituximab in patients with CIDP and anti-paranodal protein antibodies. Secondary outcome measures include grip strength, manual muscle testing sum scores, results of nerve conduction studies, and other functional scales.

Results: We plan to enroll 25 cases for the full analysis set. Recruitment is ongoing, with 14 patients enrolled as of January 2020. Enrollment will close in September 2020, and the study is planned to end in December 2021.

Conclusions: This randomized controlled trial will determine if rituximab is safe and effective in patients with anti-paranodal antibodies. An open-label study will provide additional data on the effects of rituximab in patients with antibody-negative CIDP.

The results of the RECIPE trial are expected to provide evidence for the positioning of rituximab as a pathogenesis-based therapeutic for refractory CIDP.

Trial Registration: ClinicalTrials.gov NCT03864185, https://clinicaltrials.gov/ct2/show/NCT03864185 ; The Japan Registry of Clinical Trials jRCT2041180037, https://jrct.niph.go.jp/en-latest-detail/jRCT2041180037

International Registered Report Identifier (IRRID): DERR1-10.2196/17117

(JMIR Res Protoc 2020;9(4):e17117) doi:10.2196/17117

KEYWORDS

chronic inflammatory demyelinating polyradiculoneuropathy; rituximab; immunoglobulin G4 autoantibodies; clinical study; partially randomized controlled trial

Introduction

Chronic inflammatory demyelinating polyradiculoneuropathy (CIDP) is an acquired demyelinating peripheral neuropathy characterized by weakness and sensory disturbance in the extremities that develops over a period of at least 2 months. The current diagnostic criteria for CIDP were published by the European Federation of Neurological Societies/Peripheral Nerve Society (EFNS/PNS) and include clinical and electrophysiological findings with supporting features that strengthen the diagnosis [1-4]. The global prevalence of CIDP diagnosed using these criteria is around 3 per 100,000 population, and it is considered an orphan disease. However, the etiology of CIDP has not been fully clarified, and a biomarker reflecting its pathogenesis is lacking, so the exact prevalence remains uncertain [5-10]. One of the reasons why the etiology remains unclear might be the diagnostic vulnerability of CIDP. Although the EFNS/PNS criteria identify the typical form and 5 atypical forms by phenotype, there is no biomarker that explains each phenotype. Moreover, although a wide range of cellular and humoral autoimmunity factors are associated with CIDP onset, there is no specific immune mechanism that corresponds to each atypical CIDP subtype, although macrophages are assumed to be the main effector in typical CIDP.

Autoantibodies targeting myelin components, inflammatory cytokines, and the complement pathway are assumed to be the humoral factors modifying immune status in CIDP. Although many autoantibodies against various myelin antigens, such as P0, PMP22, Cx32, beta-tubulin, LM1, and sulfatide, have been investigated, none has been implicated in the pathogenesis of CIDP [11,12]. Autoantibodies that target contactin-1 (CNTN1), neurofascin-155 (NF155), and contactin-associated protein 1 (CASPR 1) have been identified in patients with slowly progressing CIDP phenotypes. The anti-CNTN1 antibody has been identified in only a small percentage of patients with CIDP, the anti-NF155 antibody has been identified in approximately 10%, and the anti-CASPR 1 antibody appears to be very rare [13-15]. Most anti-NF155 antibody–positive patients tend to develop tremor and sensory ataxia with distal weakness and muscle atrophy. They also have extremely high cerebrospinal fluid protein levels (usually >200 mg/dL), and gadolinium-enhanced magnetic resonance imaging shows severe swelling of the nerve roots or plexus. CIDP with anti-NF155 and anti-CNTN1 antibodies have several clinically similar characteristics. However, compared with anti-NF155–positive patients, anti-CNTN1–positive patients are relatively older at disease onset and have more rapid disease development. Approximately 70% of patients with these immunoglobulin (Ig)G4 autoantibodies are resistant to first-line therapies, such as intravenous immunoglobulin (IVIg) [13,14,16-18]. IVIg has immunomodulatory activity via several mechanisms, including immune regulation by macrophages or antigen-presenting cells via Fc receptors and idiotypic antibodies, suppression of activating cytokines, and the complement pathway. However, it would be inappropriate to induce Igs to inhibit complement-associated pathogenesis in patients with IgG4 autoantibody–positive CIDP because IgG4 lacks a complement-binding site and has no or limited ability to activate the classical complement pathway. Moreover, IVIg has been found ineffective as initial treatment for IgG4 autoantibody–positive CIDP. Long-term use of corticosteroids is challenging because of adverse reactions, including susceptibility to infection, osteoporosis, abnormal glucose tolerance, hyperlipidemia, and, possibly, psychiatric symptoms and insomnia. Plasmapheresis is a first- or second-line therapeutic strategy for CIDP but is difficult to implement as maintenance therapy because of high costs and safety issues associated with vascular access and fluctuations in plasma circulation. Therefore, novel therapeutics that can achieve long-term suppression of pathogenic antibodies are needed for IgG4 autoantibody–positive CIDP. Rituximab, a chimeric anti-CD20 monoclonal antibody, selectively depletes B cells and might suppress production of antibodies by inhibiting their differentiation into plasma cells. Also, some case reports suggest that rituximab is effective in refractory CIDP [19-29]. However, all these reports were retrospective and uncontrolled, and there is little evidence to suggest that rituximab is effective for IgG4 autoantibody–positive CIDP [30-32]. Moreover, the efficacy of rituximab has been proven in other IgG4 diseases such as anti-muscle-specific tyrosine kinase (MuSK) myasthenia gravis [33,34].

Therefore, we planned this prospective clinical trial to investigate the efficacy of rituximab in refractory IgG4 autoantibody–positive CIDP by comparing with placebo in IgG4 autoantibody–positive patients and with rituximab in IgG4 autoantibody–negative patients as a reference. This investigator-initiated clinical trial has been named “The Evaluation of Efficacy and Safety of Rituximab (Genetical Recombination) in Refractory Chronic Inflammatory Demyelinating Polyneuropathy (CIDP) Patients with Immunoglobulin G4 (IgG4) Autoantibodies in the Exploratory

Clinical Trial" (RECIPE trial). The design of the trial was developed in consultation with and approved by the Pharmaceuticals and Medical Device Agency (PMDA), which is responsible for reviewing new pharmaceuticals in Japan.

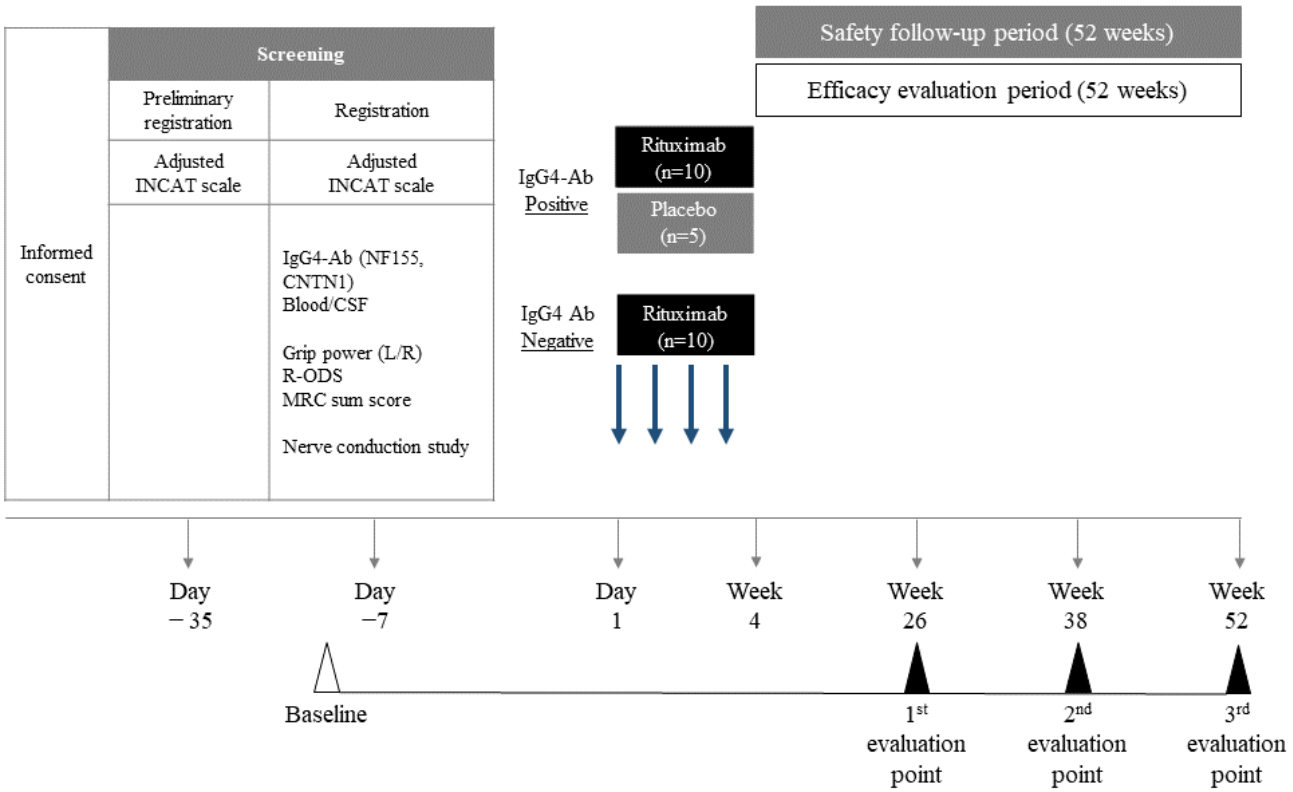
Here, we provide the detailed design of this investigator-initiated clinical trial in refractory CIDP patients with or without IgG4 autoantibodies as an exploratory study in Japan (the RECIPE trial).

Methods

Objectives and Endpoints

The primary objectives of the RECIPE trial are to evaluate the efficacy and safety of intravenous rituximab in patients with refractory CIDP according to their IgG4 autoantibody status. This trial consists of a multicenter, placebo-controlled, randomized, double-blind, parallel-group, comparative study in 15 patients with CIDP and IgG4 autoantibodies (CNTN1 or NF155) who will be allocated to a rituximab group (n=10) or a placebo group (n=5) and an open-label study in 10 patients with IgG4 autoantibody–negative CIDP. The outline of this study is shown in Figure 1.

Figure 1. Measurement schedule. CNTN!: Contactin-1; CSF: cerebrospinal fluid; IgG4-Ab: immunoglobulin G4 antibodies; INCAT: Inflammatory Neuropathy Cause and Treatment; MRC: Medical Research Council; NF155: Neurofascin-155; R-ODS: Rasch-built Overall Disability Scale.



The primary endpoint is improvement in the adjusted Inflammatory Neuropathy Cause and Treatment (INCAT) Disability Scale score at 26, 38, or 52 weeks after the start of treatment. Efficacy is based on the effectiveness rate, which is the proportion of patients who achieve the primary endpoint of ≥1-point improvement in the INCAT Disability Scale score compared with baseline. Our specific hypothesis is that the effectiveness rate will exceed a prespecified threshold of 15% in patents with IgG4 autoantibody–positive CIDP in the rituximab-treated group. The data from patients with IgG4 autoantibody–negative CIDP will serve as a reference.

Study Population

The RECIPE trial will investigate the efficacy of rituximab in patients with refractory IgG4 autoantibody–positive and IgG4 autoantibody–negative CIDP. Anti-NF155 and anti-CNTN1 autoantibody status will be confirmed by enzyme-linked immunosorbent assay as described in the literature [17,35].

Inclusion Criteria

Patients who meet the following criteria are eligible to participate:

1. CIDP diagnosed according to the modified EFNS/PNS criteria [1-4] prior to enrollment
2. Positive or negative serum IgG4 autoantibody (CNTN1 or NF155) status confirmed prior to enrollment
3. CIDP refractory to treatment with corticosteroids for 12 weeks and IVIg for 8 weeks prior to enrollment or unable to receive or continue treatment with corticosteroids and IVIg
4. Total adjusted INCAT Disability Scale scores of 2 to 8 at both preliminary enrollment and enrollment, with no

decrease in total score between preliminary enrollment and enrollment
5. Age ≥12 years
6. Able to provide voluntary written consent after receiving adequate information about the study (for children aged 12 to 15 years, consent will be obtained from an acceptable representative, and informed assent will be obtained from the patient)

Exclusion Criteria

Patients are excluded for any of the following reasons:

1. Any differential diagnosis defined in the modified EFNS/PNS diagnostic criteria [1-4]
2. Initiation or an increased dose of corticosteroids within 12 weeks before enrollment
3. Initiation or an increased dose of IVIg within 8 weeks before enrollment
4. Plasmapheresis within 8 weeks before enrollment or disease refractory to 8 weeks of plasma exchange or double-filtration plasmapheresis
5. Initiation or an increased dose of an immunosuppressant (ie, azathioprine, cyclophosphamide, cyclosporine, mycophenolate mofetil, interferon-alpha, interferon-beta, etanercept, methotrexate, mitoxantrone, alemtuzumab, cladribine, tacrolimus, or fingolimod) within 12 weeks before enrollment
6. History of hematopoietic stem cell transplantation
7. Previous treatment with rituximab
8. Participation in another clinical study within 3 months before enrollment or current participation in another study
9. Poorly controlled diabetes (hemoglobin A_{1c} ≥7%)
10. Suspected or confirmed infection requiring systemic treatment with an antimicrobial, antifungal, or antiviral agent at enrollment
11. Positive for hepatitis B surface antigen or antibody, hepatitis B core antibody, hepatitis C virus antibody, or HIV or human T-cell lymphotropic virus-1 antibody positivity at enrollment
12. Leukopenia (<2000/mm^3), neutropenia (<1000/mm^3), or lymphopenia (<500/mm^3) at enrollment
13. History of severe hypersensitivity or anaphylactic reaction to any of the ingredients in the investigational drug or murine protein-containing products
14. Severe comorbidity (eg, hepatic, renal, cardiac, lung, hematologic, or brain disease)
15. Pregnancy or potential for pregnancy, lactation, or unwillingness to use contraception during the study period
16. Deemed unsuitable for participation by an investigator or sub-investigator

Regarding item 11, patients with positive hepatitis B surface antibody or core antibody can be enrolled when a hepatitis B virus-DNA test is negative (below the limit of detection), and hepatitis B virus-DNA and aspartate/alanine transaminase levels are monitored at fixed intervals.

Registration and Randomization

The investigators obtain written informed consent from all patients who are potentially eligible for the study. After confirmation of inclusion criteria 4 to 6 and exclusion criteria 1, 6, 7, and 13 to 16, patients are preliminarily enrolled via an electronic data capture system (Viedoc™, PCG Solutions Ab. Uppsala, Sweden). At the second screening, the investigators check all the inclusion and exclusion criteria. Patients confirmed to have IgG4 autoantibody–positive CIDP are randomly assigned to the rituximab group or the placebo group at a ratio of 2:1 using the stratified permuted block method according to their adjusted INCAT Disability Scale score (2 to 4 or 5 to 8). Patients with IgG4 autoantibody–negative CIDP are automatically allocated to the rituximab group.

Investigational Treatment

Rituximab 375 mg/m^2 or placebo is administered intravenously once weekly for 4 weeks. This is the same dosage as used in previous studies involving CIDP patients treated with rituximab [20-25,27,30,31]. This dosage is also the same as that used in a phase I/II trial that showed peripheral B cells (CD20-positive) were adequately decreased in patients with B-cell non-Hodgkin's lymphoma or nephrotic syndrome. The time taken for these cells to return to their baseline level was at least 5-7 months in that study.

Hematopoietic stem cell transplantation, plasma exchange therapy (simple plasma exchange or double membrane filtration), and any drug that could affect immune status is prohibited during the trial in view of their potential to affect the efficacy of rituximab. For the same reason, starting or increasing the dose of corticosteroid, IVIg, or immunosuppressant therapy is not permitted for the duration of the study.

Primary Endpoint

All participants are followed for 52 weeks after the first dose of investigational medication (Table 1). All data are collected via an electronic data capture system and checked according to the data management and monitoring plan. Adjusted INCAT Disability Scale score is used to evaluate lower (gait) and upper extremity disorders; this tool has been used as an efficacy endpoint in clinical investigations of IVIg as a treatment for CIDP [36,37]. A change of ≥1 point in any of the items on this scale is considered clinically significant in terms of the ability to perform activities of daily living. Therefore, we selected a change of ≥1 point in the adjusted INCAT Disability Scale score as the primary endpoint. The primary analysis will compare the adjusted INCAT Disability Scale scores recorded before treatment (at enrollment) and those recorded at week 26, 38, or 52.

Table 1. Protocol for data collection from each patient enrolled in the RECIPE trial.

Visit	Screening								
	Preliminary registration	Registration	Administration	Weeks 1, 2, and 3	Week 4	Week 12	Week 26	Wek 38	Week 52 or withdrawal
Allowance (days)	up to –35	–34 to –7	–3	–3	±3	±14	±14	±14	±28
Informed consent									
Administration of investigational drug			x	x					
Hospitalization			x	x	x				
Basic information (eg, birth date, sex, body weight)	x	x							
Eligibility	x	x							
Pregnancy test	x	x							
Vital signs[a]		x	x	x	x	x	x	x	x
Oxygen saturation			x[b]	x[b]	x				
Hematology[c]		x	x	x	x	x	x	x	x
Blood chemistry[d]		x	x	x	x	x	x	x	x
Other blood tests[e]		x							
Urinalysis[f]		x	x	x	x	x	x	x	x
Chest X-ray		x				x	x		x
Adjusted INCAT[g] Disability Scale	x[h]	x[h]			x	x	x	x	x
Grip power (vigorimeter) Left/Right		x			x	x	x	x	x
Rasch-built Overall Disability Scale		x			x	x	x	x	x
MRC-SS[i]		x			x	x	x	x	x
Nerve conduction study		x			x		x	x	x
CSF[j] protein		x					x	x	x
Serum autoantibody[k]	x[l]						x	x	x
Serum neurofilament light		x					x	x	x
B cell/T cell (whole blood, %)[m]			x		x	x	x	x	x
HACA[m,n]			x				x		x
Pharmacokinetics[m]		x	x[o]	x[o]	x	x	x	x	x
Concomitant medication	x	x	x	x	x	x	x	x	x
Adverse event(s)			x	x	x	x	x	x	x

[a]Blood pressure, pulse rate, and body temperature.

[b]Measured within 30 minutes before the start of dose administration; immediately before a change in dosing rate, dose interruption, or dose reduction; before dose administration is resumed; and within 10 minutes and 1 hour after the end dose administration.

[c]Red blood cell count, hemoglobin, hematocrit, white blood cell count, differential white blood count (basophils, eosinophils, neutrophils, lymphocytes, and monocytes), and platelet count.

[d]Blood urea nitrogen, creatinine, lactate dehydrogenase, aspartate/alkaline transaminase, alkaline phosphatase, gamma glutamyl transpeptidase, total bilirubin, direct bilirubin, creatine kinase, and C-reactive protein.

[e]Glycated hemoglobin, hepatitis B surface antigen, hepatitis B surface antibody, hepatitis B virus-DNA if needed, hepatitis B core antibody, hepatitis C virus antibody, HIV antibody, and human T-cell lymphotropic virus-1.

[f]Urine protein, urine occult blood, urine glucose, and pH.

[g]INCAT: Inflammatory Neuropathy Cause and Treatment.

[h]If more than 28 days has passed between provisional enrollment and enrollment.

[i]MRC-SS: Medical Research Council Sum-Score.

[j]CSF: cerebrospinal fluid.

[k]Further measurement is not required if not detected at screening.

[l]Measured only once during the screening period.

[m]Test results are not disclosed during the study period.

[n]HACA: human anti-chimeric antibody.

[o]Blood is collected within 15 minutes before the start and after the end of dose administration.

Secondary, Exploratory, and Safety Endpoints

Secondary endpoints include grip strength, as measured using a vigorimeter (left/right); the Rasch-built Overall Disability Scale; Medical Research Council sum score; median, ulnar, tibial, and peroneal motor nerve conduction studies, including motor nerve conduction velocity and distal and proximal latency, amplitude, and duration of the compound muscle action potential after distal and proximal stimulation; protein levels in the cerebrospinal fluid; B cell (CD19-positive and CD20-positive) counts and T-cell (CD3-positive, CD4-positive, and CD8-positive) counts; expression of human anti-chimeric antibodies; and serum rituximab level.

The exploratory endpoints include serum antibody titers of IgG4, serum antibody titers of anti-CNTN1 and anti-NF155 IgG subclasses, and serum neurofilament level.

The safety endpoints include adverse events and vital signs and laboratory tests.

Sample Size

The trial includes patients with CIDP that is refractory to both IVIg and corticosteroids or patients who are unable to receive or continue these therapies. Most eligible patients would be unlikely to achieve an improvement of ≥1 point on the adjusted INCAT Disability Scale even if conventional therapies were continued. An immunosuppressant (eg, cyclophosphamide, azathioprine, or cyclosporine) is generally considered to be an effective second-line treatment in about 30% of patients with refractory CIDP [19]. However, no superiority has been demonstrated for any specific agent, and no immunosuppressant is approved for CIDP in Japan. Currently, no feasible treatment approach can be recommended for patients enrolled in this study. Therefore, the threshold effectiveness rate for the proportion of patients achieving the primary endpoint of ≥1 point improvement in the adjusted INCAT Disability Scale score was conservatively assumed to be 15%, and the clinically expected effectiveness rate was assumed to be 60% in patients with IgG4 autoantibody–positive CIDP in the rituximab group. Using two-sided testing with $P<0.05$ regarded as statistically significant, we estimated that 8 cases would be needed to achieve 80% power. Accordingly, 10 patients with IgG4 autoantibody–positive CIDP will be enrolled for the rituximab group, allowing for a 20% dropout rate.

The PMDA suggested inclusion of a placebo control group for the patients who are IgG4 autoantibody–positive. We agreed to include a placebo group of 5 patients for reference purposes rather than for strict statistical comparison, given the rarity of CIDP and feasibility considerations. The placebo group has been designated as a reference group to confirm that the estimated values do not deviate markedly from the defined thresholds. Patients with IgG4 autoantibody–negative CIDP are included in another reference group to compare their response to rituximab treatment with that of patients with IgG4 autoantibody–positive CIDP. This reference group will consist of 10 patients, which is the same as the number of patients with IgG4 autoantibody–positive CIDP.

Statistical Analysis

All statistical analyses will be performed on an intention-to-treat basis. The primary analysis will compare the adjusted INCAT Disability Scale scores recorded before treatment (at enrollment) with those recorded at 26, 38, or 52 weeks after the start of treatment. We will calculate the proportion and 95% CI of IgG4 autoantibody–positive patients in the rituximab group who achieve an improvement of ≥1 point at any of evaluation point after 26 weeks, from baseline. The 95% CI will be calculated using the Clopper-Pearson method. Continuous variables will be analyzed using descriptive statistics, and categorical variables will be calculated as the frequency and proportion. All statistical analyses will be performed using SAS version 9.4 (SAS Institute Inc, Cary, NC). All statistical tests will be two-sided. *P* values <0.05 will be considered statistically significant.

Monitoring and Auditing

Monitoring and auditing will include systematic independent examination according to the study protocol, applicable regulatory requirements, and standard operating procedures.

Results

Ethics Approval, Trial Registration, and Current Enrollment Status

The study protocol complies with the Declaration of Helsinki [38] and the Pharmaceutical Affairs Act in Japan. This protocol was also approved by the institutional review boards at the following sites: Nagoya University Hospital (No. 302010), Chiba University Hospital (No. 030033), Yamaguchi University Hospital (No. 201901), and Kyushu University Hospital (No. 2018312).

The necessary information for the RECIPE trial has been uploaded to ClinicalTrials.gov (NCT03864185, registered March 6, 2019) and the Japan Registry of Clinical Trials (jRC2041180037, registered January 31, 2019).

The first patient completed registration in April 2019 and received an investigational treatment in May 2019. Recruitment of patients for the RECIPE trial is ongoing at the four participating hospitals. As of January 2020, 14 cases have been enrolled. The targeted accrual is 25 cases for the full analysis set. Enrollment will close in September 2020, and the study is scheduled to end in December 2021.

Discussion

Overview

There have been some reports on the epidemiology of IgG4 antibody–positive CIDP [14,16,17]. This phenotypic subtype presents as subacute or slowly progressive disease. The initial disability tends to be distal acquired demyelinating symmetric neuropathy, which sometimes progresses to the typical CIDP phenotype. IgG4 antibody–positive CIDP is also characterized by gait disturbance with sensory ataxia and fine tremor of the hands. The protein level in cerebrospinal fluid is markedly higher in IgG4 antibody–positive CIDP than in typical CIDP. Furthermore, a relatively large proportion of patients with antibody–positive CIDP present at a younger age. However, they are resistant to conventional therapies, such as IVIg and corticosteroids. Therefore, new therapies that can inhibit production of pathogenic antibodies in the long term are essential.

Rituximab can be expected to be effective in cases of refractory CIDP [19-32]. Some of the phenotypic characteristics, for example onset time, are different from those in anti-CNTN1–positive patients and anti-NF155 patients; however, patients with IgG4 autoantibody–positive CIDP have several clinically common characteristics. Given its mechanism of action, rituximab is likely to be effective in patients with IgG4 autoantibody (anti-CNTN1 and anti-NF155)–positive CIDP. Accordingly, we are planning to develop rituximab for use in these patients and present here the protocol for the exploratory study intended to secure approval of rituximab in Japan. We have attended a consultation meeting with the PMDA, which has agreed to the development strategy and initial design of the RECIPE trial.

Conclusion

This article has described the design and protocol being used in the RECIPE trial. We outlined the characteristics of patients with IgG4 autoantibody–positive CIDP and discussed some critical considerations for these patients. The RECIPE trial is the first randomized controlled trial of rituximab for IgG4 autoantibody–positive CIDP. It is anticipated that the results will lead to a pathogenesis-oriented therapeutic strategy that can target specific phenotypes of CIDP and confirm the efficacy of rituximab in refractory IgG4 autoantibody–negative CIDP.

Acknowledgments

We thank the staff at all the participating research facilities (Nagoya University Hospital, Chiba University Hospital, Yamaguchi University Hospital, Kyushu University Hospital) for assisting with the recruitment of patients for this trial. We also thank the staff of the National Defense Medical College Hospital for analysis of NF155 and CNTN1 autoantibody status and IgG subclass. We are grateful to Zenyaku Kogyo Co, Ltd for providing the investigational drug (rituximab) and placebo, collection of safety information, keycode registration, transportation of samples, and measurement of B-cell, T-cell, human anti-chimeric antibody, and serum rituximab levels. We thank the Japan Agency for Medical Research and Development for its financial support of this research (grant number JP19lk0201080), which is covering almost all the operational costs of this trial.

Conflicts of Interest

None declared.

References

1. Joint Task Force of the EFNS and the PNS. European Federation of Neurological Societies/Peripheral Nerve Society Guideline on management of chronic inflammatory demyelinating polyradiculoneuropathy: report of a joint task force of the European Federation of Neurological Societies and the Peripheral Nerve Society--First Revision. J Peripher Nerv Syst 2010 Mar;15(1):1-9. [doi: 10.1111/j.1529-8027.2010.00245.x] [Medline: 20433600]
2. Van den Bergh PYK, Hadden RDM, Bouche P, Cornblath DR, Hahn A, Illa I, European Federation of Neurological Societies, Peripheral Nerve Society. European Federation of Neurological Societies/Peripheral Nerve Society guideline on management of chronic inflammatory demyelinating polyradiculoneuropathy: report of a joint task force of the European Federation of Neurological Societies and the Peripheral Nerve Society - first revision. Eur J Neurol 2010 Mar;17(3):356-363. [doi: 10.1111/j.1468-1331.2009.02930.x] [Medline: 20456730]
3. Joint Task Force of the EFNS and the PNS. Erratum of European Federation of Neurological Societies/Peripheral Nerve Society Guideline on management of chronic inflammatory demyelinating polyradiculoneuropathy: report of a joint task force of the European Federation of Neurological Societies and the Peripheral Nerve Society – First Revision. J Peripher Nerv Syst 2010;15:373 [FREE Full text] [doi: 10.1111/j.1529-8027.2010.00291.x]
4. Van den Bergh PYK, Hadden RDM, Bouche P, Cornblath DR, Hahn A, Illa I, European Federation of Neurological Societies, Peripheral Nerve Society. Corrigendum of European Federation of Neurological Societies/Peripheral Nerve Society Guideline on management of chronic inflammatory demyelinating polyradiculoneuropathy: Report of a joint task force of the European Federation of Neurological Societies and the Peripheral Nerve Society — First Revision. Eur J Neurol 2011 Apr;18(5):796 [FREE Full text] [doi: 10.1111/j.1468-1331.2011.03403.x]

5. Laughlin RS, Dyck PJ, Melton LJ, Leibson C, Ransom J, Dyck PJB. Incidence and prevalence of CIDP and the association of diabetes mellitus. Neurology 2009 Jul 07;73(1):39-45 [FREE Full text] [doi: 10.1212/WNL.0b013e3181aaea47] [Medline: 19564582]
6. Rajabally YA, Simpson BS, Beri S, Bankart J, Gosalakkal JA. Epidemiologic variability of chronic inflammatory demyelinating polyneuropathy with different diagnostic criteria: study of a UK population. Muscle Nerve 2009 Apr;39(4):432-438. [doi: 10.1002/mus.21206] [Medline: 19260065]
7. Lunn MP, Manji H, Choudhary PP, Hughes RA, Thomas PK. Chronic inflammatory demyelinating polyradiculoneuropathy: a prevalence study in south east England. J Neurol Neurosurg Psychiatry 1999 May;66(5):677-680 [FREE Full text] [doi: 10.1136/jnnp.66.5.677] [Medline: 10209187]
8. Chiò A, Cocito D, Bottacchi E, Buffa C, Leone M, Plano F, PARCIDP. Idiopathic chronic inflammatory demyelinating polyneuropathy: an epidemiological study in Italy. J Neurol Neurosurg Psychiatry 2007 Dec;78(12):1349-1353 [FREE Full text] [doi: 10.1136/jnnp.2007.114868] [Medline: 17494979]
9. Mygland A, Monstad P. Chronic polyneuropathies in Vest-Agder, Norway. Eur J Neurol 2001 Mar;8(2):157-165. [doi: 10.1046/j.1468-1331.2001.00187.x] [Medline: 11284994]
10. McLeod JG, Pollard JD, Macaskill P, Mohamed A, Spring P, Khurana V. Prevalence of chronic inflammatory demyelinating polyneuropathy in New South Wales, Australia. Ann Neurol 1999 Dec;46(6):910-913. [Medline: 10589544]
11. Meléndez-Vásquez C, Redford J, Choudhary PP, Gray IA, Maitland P, Gregson NA, et al. Immunological investigation of chronic inflammatory demyelinating polyradiculoneuropathy. J Neuroimmunol 1997 Mar;73(1-2):124-134. [doi: 10.1016/s0165-5728(96)00189-0] [Medline: 9058768]
12. Kuwahara M, Suzuki S, Takada K, Kusunoki S. Antibodies to LM1 and LM1-containing ganglioside complexes in Guillain-Barré syndrome and chronic inflammatory demyelinating polyneuropathy. J Neuroimmunol 2011 Oct 28;239(1-2):87-90. [doi: 10.1016/j.jneuroim.2011.08.016] [Medline: 21914557]
13. Querol L, Siles AM, Alba-Rovira R, Jáuregui A, Devaux J, Faivre-Sarrailh C, et al. Antibodies against peripheral nerve antigens in chronic inflammatory demyelinating polyradiculoneuropathy. Sci Rep 2017 Oct 31;7(1):14411 [FREE Full text] [doi: 10.1038/s41598-017-14853-4] [Medline: 29089585]
14. Devaux JJ, Miura Y, Fukami Y, Inoue T, Manso C, Belghazi M, et al. Neurofascin-155 IgG4 in chronic inflammatory demyelinating polyneuropathy. Neurology 2016 Mar 01;86(9):800-807 [FREE Full text] [doi: 10.1212/WNL.0000000000002418] [Medline: 26843559]
15. Miura Y, Devaux JJ, Fukami Y, Manso C, Belghazi M, Wong AHY, CNTN1-CIDP Study Group. Contactin 1 IgG4 associates to chronic inflammatory demyelinating polyneuropathy with sensory ataxia. Brain 2015 Jun;138(Pt 6):1484-1491 [FREE Full text] [doi: 10.1093/brain/awv054] [Medline: 25808373]
16. Ogata H, Yamasaki R, Hiwatashi A, Oka N, Kawamura N, Matsuse D, et al. Characterization of IgG4 anti-neurofascin 155 antibody-positive polyneuropathy. Ann Clin Transl Neurol 2015 Oct;2(10):960-971 [FREE Full text] [doi: 10.1002/acn3.248] [Medline: 26478896]
17. Kadoya M, Kaida K, Koike H, Takazaki H, Ogata H, Moriguchi K, et al. IgG4 anti-neurofascin155 antibodies in chronic inflammatory demyelinating polyradiculoneuropathy: Clinical significance and diagnostic utility of a conventional assay. J Neuroimmunol 2016 Dec 15;301:16-22. [doi: 10.1016/j.jneuroim.2016.10.013] [Medline: 27852440]
18. Querol L, Nogales-Gadea G, Rojas-Garcia R, Diaz-Manera J, Pardo J, Ortega-Moreno A, et al. Neurofascin IgG4 antibodies in CIDP associate with disabling tremor and poor response to IVIg. Neurology 2014 Mar 11;82(10):879-886 [FREE Full text] [doi: 10.1212/WNL.0000000000000205] [Medline: 24523485]
19. Cocito D, Grimaldi S, Paolasso I, Falcone Y, Antonini G, Benedetti L, Italian Network for CIDP Register. Immunosuppressive treatment in refractory chronic inflammatory demyelinating polyradiculoneuropathy. A nationwide retrospective analysis. Eur J Neurol 2011 Dec;18(12):1417-1421. [doi: 10.1111/j.1468-1331.2011.03495.x] [Medline: 21819489]
20. Levine TD, Pestronk A. IgM antibody-related polyneuropathies: B-cell depletion chemotherapy using Rituximab. Neurology 1999 May 12;52(8):1701-1704. [doi: 10.1212/wnl.52.8.1701] [Medline: 10331706]
21. Knecht H, Baumberger M, Tobòn A, Steck A. Sustained remission of CIDP associated with Evans syndrome. Neurology 2004 Aug 24;63(4):730-732. [doi: 10.1212/01.wnl.0000134606.50529.c7] [Medline: 15326255]
22. Gono T, Matsuda M, Shimojima Y, Ishii W, Yamamoto K, Morita H, et al. Rituximab therapy in chronic inflammatory demyelinating polyradiculoneuropathy with anti-SGPG IgM antibody. J Clin Neurosci 2006 Jul;13(6):683-687. [doi: 10.1016/j.jocn.2005.09.008] [Medline: 16814550]
23. Münch C, Anagnostou P, Meyer R, Haas J. Rituximab in chronic inflammatory demyelinating polyneuropathy associated with diabetes mellitus. J Neurol Sci 2007 May 15;256(1-2):100-102. [doi: 10.1016/j.jns.2007.02.027] [Medline: 17382963]
24. Gorson KC, Natarajan N, Ropper AH, Weinstein R. Rituximab treatment in patients with IVIg-dependent immune polyneuropathy: a prospective pilot trial. Muscle Nerve 2007 Jan;35(1):66-69. [doi: 10.1002/mus.20664] [Medline: 16967492]
25. Benedetti L, Briani C, Franciotta D, Fazio R, Paolasso I, Comi C, et al. Rituximab in patients with chronic inflammatory demyelinating polyradiculoneuropathy: a report of 13 cases and review of the literature. J Neurol Neurosurg Psychiatry 2011 Mar;82(3):306-308. [doi: 10.1136/jnnp.2009.188912] [Medline: 20639381]
26. Sadnicka A, Reilly MM, Mummery C, Brandner S, Hirsch N, Lunn MPT. Rituximab in the treatment of three coexistent neurological autoimmune diseases: chronic inflammatory demyelinating polyradiculoneuropathy, Morvan syndrome and

myasthenia gravis. J Neurol Neurosurg Psychiatry 2011 Feb;82(2):230-232. [doi: 10.1136/jnnp.2009.174888] [Medline: 20462915]
27. D'Amico A, Catteruccia M, De Benedetti F, Vivarelli M, Colucci M, Cascioli S, et al. Rituximab in a childhood-onset idiopathic refractory chronic inflammatory demyelinating polyneuropathy. Eur J Paediatr Neurol 2012 May;16(3):301-303. [doi: 10.1016/j.ejpn.2011.08.002] [Medline: 21903431]
28. Ware TL, Kornberg AJ, Rodriguez-Casero MV, Ryan MM. Childhood chronic inflammatory demyelinating polyneuropathy: an overview of 10 cases in the modern era. J Child Neurol 2014 Jan;29(1):43-48. [doi: 10.1177/0883073812471719] [Medline: 23364655]
29. Velardo D, Riva N, Del Carro U, Bianchi F, Comi G, Fazio R. Rituximab in refractory chronic inflammatory demyelinating polyradiculoneuropathy: report of four cases. J Neurol 2017 May;264(5):1011-1014. [doi: 10.1007/s00415-017-8462-7] [Medline: 28337614]
30. Querol L, Rojas-García R, Diaz-Manera J, Barcena J, Pardo J, Ortega-Moreno A, et al. Rituximab in treatment-resistant CIDP with antibodies against paranodal proteins. Neurol Neuroimmunol Neuroinflamm 2015 Sep 3;2(5):e149 [FREE Full text] [doi: 10.1212/NXI.0000000000000149] [Medline: 26401517]
31. Delmont E, Manso C, Querol L, Cortese A, Berardinelli A, Lozza A, et al. Autoantibodies to nodal isoforms of neurofascin in chronic inflammatory demyelinating polyneuropathy. Brain 2017 Jul 01;140(7):1851-1858. [doi: 10.1093/brain/awx124] [Medline: 28575198]
32. Painous C, López-Pérez M, Illa I, Querol L. Head and voice tremor improving with immunotherapy in an anti-NF155 positive CIDP patient. Ann Clin Transl Neurol 2018 May 7;5(4):499-501 [FREE Full text] [doi: 10.1002/acn3.539] [Medline: 29687027]
33. Hehir MK, Hobson-Webb LD, Benatar M, Barnett C, Silvestri NJ, Howard JF, et al. Rituximab as treatment for anti-MuSK myasthenia gravis: Multicenter blinded prospective review. Neurology 2017 Sep 05;89(10):1069-1077. [doi: 10.1212/WNL.0000000000004341] [Medline: 28801338]
34. Díaz-Manera J, Martínez-Hernández E, Querol L, Klooster R, Rojas-García R, Suárez-Calvet X, et al. Long-lasting treatment effect of rituximab in MuSK myasthenia. Neurology 2012 Jan 17;78(3):189-193. [doi: 10.1212/WNL.0b013e3182407982] [Medline: 22218276]
35. Koike H, Kadoya M, Kaida K, Ikeda S, Kawagashira Y, Iijima M, et al. Paranodal dissection in chronic inflammatory demyelinating polyneuropathy with anti-neurofascin-155 and anti-contactin-1 antibodies. J Neurol Neurosurg Psychiatry 2017 Jun;88(6):465-473. [doi: 10.1136/jnnp-2016-314895] [Medline: 28073817]
36. van Schaik IN, Bril V, van Geloven N, Hartung H, Lewis RA, Sobue G, PATH study group. Subcutaneous immunoglobulin for maintenance treatment in chronic inflammatory demyelinating polyneuropathy (PATH): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Neurol 2018 Jan;17(1):35-46. [doi: 10.1016/S1474-4422(17)30378-2] [Medline: 29122523]
37. Léger J, De Bleecker JL, Sommer C, Robberecht W, Saarela M, Kamienowski J, PRIMA study investigators. Efficacy and safety of Privigen(®) in patients with chronic inflammatory demyelinating polyneuropathy: results of a prospective, single-arm, open-label Phase III study (the PRIMA study). J Peripher Nerv Syst 2013 Jun;18(2):130-140 [FREE Full text] [doi: 10.1111/jns5.12017] [Medline: 23781960]
38. The World Medical Association. https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/. 2013 Oct. Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2020-02-09]

Abbreviations

CASPR 1: contactin-associated protein 1.
CIDP: chronic inflammatory demyelinating polyradiculoneuropathy.
CNTN1: contactin-1.
EFNS/PNS: European Federation of Neurological Societies/Peripheral Nerve Society.
Ig: immunoglobulin.
INCAT: Inflammatory Neuropathy Cause and Treatment.
IVIg: intravenous immunoglobulin.
MuSK: muscle-specific tyrosine kinase.
NF155: neurofascin-155.
PMDA: Pharmaceuticals and Medical Device Agency.

Edited by G Eysenbach; submitted 20.11.19; peer-reviewed by R Lewis, R Oki; comments to author 14.01.20; revised version received 27.01.20; accepted 27.01.20; published 01.04.20.

Please cite as:
Shimizu S, Iijima M, Fukami Y, Tamura N, Nakatochi M, Ando M, Nishi R, Koike H, Kaida K, Koga M, Kanda T, Ogata H, Kira JI, Mori M, Kuwabara S, Katsuno M
Efficacy and Safety of Rituximab in Refractory CIDP With or Without IgG4 Autoantibodies (RECIPE): Protocol for a Double-Blind, Randomized, Placebo-Controlled Clinical Trial
JMIR Res Protoc 2020;9(4):e17117
URL: https://www.researchprotocols.org/2020/4/e17117
doi:10.2196/17117
PMID:32234705

Protocol

Reducing Salt Intake in China with "Action on Salt China" (ASC): Protocol for Campaigns and Randomized Controlled Trials

Puhong Zhang[1,2*], PhD; Feng J He[3*], PhD; Yuan Li[1,2], PhD; Changning Li[4], MSc; Jing Wu[5], PhD; Jixiang Ma[6], PhD; Bing Zhang[7], PhD; Huijun Wang[7], PhD; Yinghua Li[4], PhD; Junhua Han[8], PhD; Rong Luo[1], MSc; Jing He[1], MSc; Xian Li[1,2], MSc; Yu Liu[9], PhD; Changqiong Wang[3], PhD; Monique Tan[3], MSc; Graham A MacGregor[3], FRCP; Xinhua Li[10], DSc

[1]The George Institute for Global Health at Peking University Health Science Center, Beijing, China

[2]Faculty of Medicine, University of New South Wales, Sydney, Australia

[3]Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine & Dentistry, Queen Mary University of London, London, United Kingdom

[4]Surveillance Department, Chinese Center for Health Education, Beijing, China

[5]The National Center for Chronic and Noncommunicable Disease Control and Prevention, The Chinese Center for Disease Control and Prevention, Beijing, China

[6]Chronic Diseases and Aging Health Management Division, The Chinese Center for Disease Control and Prevention, Beijing, China

[7]National Institute for Nutrition and Health, The Chinese Center for Disease Control and Prevention, Beijing, China

[8]Food Policy, China National Center for Food Safety Risk Assessment, Beijing, China

[9]School of Computing, Beihang University, Beijing, China

[10]Chinese Center for Disease Control and Prevention, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Xinhua Li, DSc
Chinese Center for Disease Control and Prevention
No 155 Changbai Road, Changping District
Beijing, 102206
China
Phone: 86 1058900201
Email: lixinhua@chinacdc.cn

Abstract

Background: Salt intake in China is over twice the maximum recommendation of the World Health Organization. Unlike most developed countries where salt intake is mainly derived from prepackaged foods, around 80% of the salt consumed in China is added during cooking.

Objective: Action on Salt China (ASC), initiated in 2017, aims to develop, implement, and evaluate a comprehensive and tailored salt reduction program for national scaling-up.

Methods: ASC consists of six programs working in synergy to increase salt awareness and to reduce the amount of salt used during cooking at home and in restaurants, as well as in processed foods. Since September 2018, two health campaigns on health education and processed foods have respectively started, in parallel with four open-label cluster randomized controlled trials (RCTs) in six provinces across China: (1) app-based intervention study (AIS), in which a mobile app is used to achieve and sustain salt reduction in school children and their families; (2) home cook-based intervention study (HIS), in which family cooks receive support in using less salt; (3) restaurant-based intervention study (RIS) targeting restaurant consumers, cooks, and managers; and (4) comprehensive intervention study (CIS), which is a real-world implementation and evaluation of all available interventions in the three other RCTs. To explore the barriers, facilitators, and effectiveness of delivering a comprehensive salt reduction intervention, these RCTs will last for 1 year (stage 1), followed by nationwide implementation (stage 2). In AIS, HIS, and CIS, the primary outcome of salt reduction will be evaluated by 24-hour urinary sodium excretion in 6030 participants, including 5436 adults and 594 school children around 8-9 years old. In RIS, the salt content of meals will be measured by laboratory food analysis of the 5 best-selling dishes from 192 restaurants. Secondary outcomes will include process evaluation; changes in knowledge, attitude, and practice on salt intake; and economic evaluation.

Results: All RCTs have been approved by Queen Mary Research Ethics Committee and the Institutional Review Boards of leading institutes in China. The research started in June 2017 and is expected to be completed around March 2021. The baseline investigations of the four RCTs were completed in May 2019.

Conclusions: The ASC project is progressing smoothly. The intervention packages and tailored components will be promoted for salt reduction in China, and could be adopted by other countries.

Trial Registration: Chinese Clinical Trial Registry. AIS: ChiCTR1800017553; https://tinyurl.com/vdr8rpr. HIS: ChiCTR1800016804; https://tinyurl.com/w8c7x3w. RIS: ChiCTR1800019694; https://tinyurl.com/uqkjgfw. CIS: ChiCTR1800018119; https://tinyurl.com/s3ajldw.

International Registered Report Identifier (IRRID): DERR1-10.2196/15933

(JMIR Res Protoc 2020;9(4):e15933) doi:10.2196/15933

KEYWORDS

sodium; dietary; 24-hour urinary sodium; salt reduction; randomized controlled trials; scaling-up; China

Introduction

Humans only require a very small amount of salt (ie, around 1 g/day) to maintain physiological function [1-3]. High salt intake is the major cause of raised blood pressure [4] and is the leading risk factor of total death and disability-adjusted life years in China [5]. Compelling evidence has shown that a lower salt intake is associated with a reduced risk of cardiovascular disease (CVD) and total mortality [6,7]. Salt reduction is one of the most cost-effective measures to prevent hypertension and CVD [4,8]. The World Health Organization (WHO) has recommended a 30% reduction in population salt intake by 2025, and also set a target of <5 g/day for all adults with even lower levels for children [1]. Accordingly, many developed countries have started salt reduction initiatives [9]. Salt intake has been successfully reduced in Finland and the United Kingdom, accompanied by falls in population blood pressure and CVD-related mortality [10]; however, developing countries are lagging behind.

China is the largest developing country, accounting for one fifth of the global population. Salt intake in China is among the highest in the world, with adults consuming on average above 10 g/day [11,12], which is more than twice the WHO recommended limit [1]. Approximately 80% of the salt in the Chinese diet is added by consumers during cooking [13]. With rapid and sustained urbanization, the amount of salt intake from restaurants and prepackaged foods is also increasing. A recent study showed that the major sources of salt intake for urban adults of working age are home cooking (50.1%, 40.8% of which is derived from cooking salt and 9.3% from various condiments), food prepared by restaurants (43.3%), and prepackaged food (6.6%) [14].

In most developed countries, where 80% of salt intake is derived from prepackaged food, the major strategy of salt reduction is to set salt targets; that is, to encourage food producers to gradually reduce the amount of salt used in their food products [15]. However, salt reduction is more challenging in China and many other developing countries owing to the difficulty in changing individuals' dietary behavior. It is of paramount importance to develop strategies and specific solutions to improve: (1) the environment, so that it encourages and facilitates salt reduction; (2) consumer knowledge, attitude, and practice (KAP) of eating food with reduced salt content; (3) family or restaurant cooks' knowledge and skills in reducing salt use during cooking; and (4) motivation of the food industry to reduce salt use in processed foods.

The central government of China has set a target of a 20% reduction in mean population salt intake by 2030 as one of the key components of China's health development agenda "Healthy China 2030" [16]. The government-led initiative "Healthy Lifestyle for All" has also identified salt reduction as one of the most important strategies to prevent noncommunicable diseases. Responding to the national call for salt reduction, several regional salt reduction projects have been undertaken in various regions of China as part of routine work in disease control and prevention systems. However, none of these existing programs has been properly evaluated for effectiveness and sustainability, and it is not known whether they can be rolled out across the whole country.

To overcome the above challenges, a collaboration unit called "Action on Salt China" (ASC) was established in June 2017, funded by the UK National Institute for Health Research (NIHR). ASC was built upon an existing collaboration between Queen Mary University of London in the United Kingdom, and The George Institute for Global Health (TGI) in China, as well as previous research and implementation experience on salt reduction [17]. In addition, ASC has included almost all of the key national organizations related to salt reduction, including the Chinese Center for Disease Control and Prevention (China CDC), Chinese Center for Health Education (CCHE), and China National Center for Food Safety Risk Assessment (CFSA), as well as local health and education authorities. The program aims to design several standardized, effective, and sustainable salt reduction packages targeting the major challenges in salt reduction in China, and to scale them up after appropriate evaluation.

ASC is running two national health campaigns and four randomized controlled trials (RCTs) testing interventions on major sources of salt intake. Although the protocol and results of each RCT will be published separately, it is worthwhile to report the overall design of ASC so as to improve public understanding of its rationale and design as a whole. With this aim, this paper introduces ASC's overall goals and strategies, governance, proposed solutions to main challenges in salt

reduction, uniformed design of intervention packages and evaluations, plan for scaling-up, as well as other ancillary work.

Methods

Goals and Objectives

The goal of ASC is to reduce salt intake by 15% by 2021 in the six target provinces in China. The specific objectives are to reduce salt intake by at least 1 g/day (about 17 mmol/day) at home, and to reduce salt use by at least 0.5 g (about 8.5 mmol) per 100 g in restaurant dishes. To reach these goals, several programs targeting the major sources of salt intake in China have been developed and are being implemented.

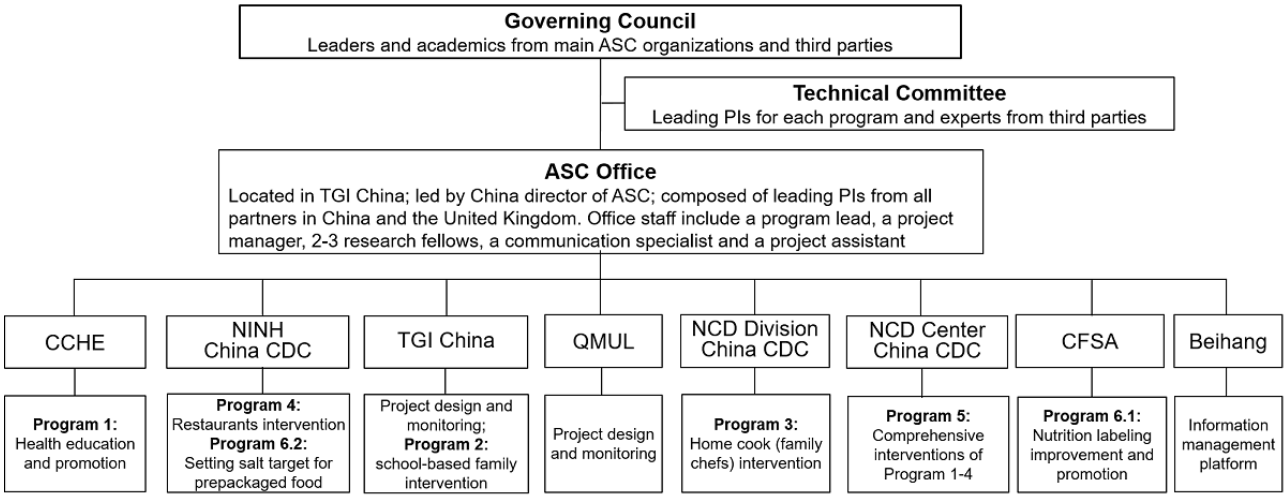
Strategy and Overall Design

To achieve these goals and objectives, the ASC project developed six programs targeting the low health literacy related to salt reduction and the three sources of salt intake in China: home cooking, restaurant foods, and prepackaged foods. Program 1 is a salt reduction education campaign, which will form the basis of all the other programs. Program 2, app-based intervention study (AIS), is a primary school-based program delivering salt reduction activities for the homes of school children, based on the assumption that parents and grandparents are more likely to change their habit of high salt intake for the sake of the health of their children and grandchildren [18,19]. Program 3, home cook-based intervention study (HIS), will establish a community-level training and support system to help family cooks reduce salt use in home cooking. Program 4, restaurant-based intervention study (RIS), aims to create an environment in restaurants that is conducive to consumers opting for reduced-salt dishes, and to train cooks to use less salt while cooking. Program 5, comprehensive intervention study (CIS), simulates the real-world implementation of all of the intervention packages or components developed in Programs 1-4 with the purpose of identifying barriers and facilitators when scaling up. Program 6 is a two-fold campaign with the aims of educating and supporting consumers choosing less salted prepackaged foods (Program 6.1), and convincing food manufacturers to reformulate their food products by gradually reducing the amount of salt added (Program 6.2).

Governance of the Action on Salt China Unit and Programs

The governance and management structure of the ASC unit and programs are illustrated in Figure 1. The governing council consists of 10 members, 5 of whom are the top leaders from the 5 member organizations of the ASC unit. The other 5 members are academics from independent organizations, including the National Preventive Medicine Association (2 members with Chairman), China Nutrition Society, Chinese Hypertension League, and Peking University Health Science Center. The governing council provides guidance by (1) reviewing ASC's goals and strategies, (2) setting the tone for cooperation and communication, and (3) evaluating ASC's overall performance and achievement every year. The technical committee is composed of the leading principal investigators of all programs and independent experts in salt reduction. The committee members will work together to finalize documents (eg, research protocols and education materials), provide support in program implementation, and act as the ultimate decision maker in handling any practical issues. Supervised and supported by the governing council and technical committee, the ASC office, located in TGI China, will be responsible for coordinating all of the partners to implement the proposed programs, and to ensure the smooth running of the whole unit and its programs to achieve quality outputs in a timely manner.

Figure 1. Governance and specific programs of Action on Salt China (ASC) program. PI: principal investigator; China CDC: China Centers for Disease Control; CCHE: Chinese Center for Health Education; NINH: National Institute for Nutrition and Health, China CDC; TGI China: The George Institute for Global Health China; QMUL: Queen Mary University of London; NCD Division: Division of NCD Control and Community Health, China CDC; NCD Center: National Center for Chronic and Non-communicable Disease Control and Prevention, China CDC; CFSA: China National Center for Food Safety Risk Assessment; Beihang: Beihang University (Or Beijing University of Aeronautics and Astronautics).



Program Implementation

A two-stage strategy is adopted to ensure that the proposed salt reduction programs are well prepared, evaluated, and tailored. At stage 1 (first 2 years), the intervention packages in each program will be developed and evaluated, before being scaled-up and integrated as policies at stage 2.

At stage 1, the different intervention packages of Programs 2-5 will be evaluated with open-label cluster RCTs in various settings. The four RCTs will be conducted collaboratively under the same governance and time frame. Evaluation will be performed using standardized questionnaires, physical examinations, laboratory tests, and data collection systems that are as identical as possible so that the data collected across the RCTs can be pooled together for overall salt intake estimation and overall effectiveness evaluation.

The key features of the six programs of ASC at stage 1, including the theories and key intervention components of the four RCTs (Programs 2-5 at stage 1), are summarized in Table 1. Separate protocols for each of the four RCTs describing the study setting and participants, randomization, intervention, sample size calculation, outcomes, and data collection and analysis will be published before the end of stage 1. These results (especially those related to effectiveness and process evaluation) will also be published separately after the RCTs are completed.

The study participants are grade 3 primary school students (8-9 years old) and their parents/grandparents (1 student and 2 adults for each family) in AIS, home cooks and their family members (1 home cook and one other adult member for each family) in HIS, and adults (1 adult from each participating family) in CIS. Although the participant recruitment may vary among AIS, HIS, and CIS, in all cases, these will be local residents with no plan to move out of the city or village within 24 months, and agree to participate in the studies. The exclusion criteria are (1) pregnant women and those actively lactating; (2) individuals who are currently participating in any other clinical trials; (3) those with severe psychiatric or physical diseases that might impact intervention and follow-up; (4) individuals who are unable or not suitable to collect 24-hour urine due to the following conditions: aconuresis; acute/chronic urinary tract infection, vaginal infection and perianal infection; acute hemorrhagic diseases in the urinary tract, vagina, and digestive tract; and severe vomiting and diarrheic symptoms. In RIS, the study subjects are restaurants with dish salt content as the primary outcome, which will be evaluated using the average sodium content of the 5 best-selling dishes.

At the end of stage 1, the education materials and the effective intervention packages or components will be combined as a scale-up intervention package on salt reduction (SIPS) for broad use at stage 2. A final evaluation to assess the impact of the scaling-up and the lasting effectiveness of the intervention packages will be carried out at the end of 1 year of scale-up. The SIPS will subsequently be further promoted over a larger scale across China using existing platforms and resources such as China's Healthy Lifestyle for All Initiative of the China CDC and the Chinese Center for Health Education.

To avoid contamination in control groups, all RCTs are being conducted in different counties or districts of the six provinces: Heilongjiang, Hebei, Hunan, Jiangxi, Sichuan, and Qinghai, which cover the north, south, central, east, and west part of China (Figure 2). HIS, RIS, and CIS are carried out in the above six provinces, whereas AIS is conducted in only three provinces: Hebei (north), Sichuan (central), and Hunan (south).

To facilitate implementation, ASCloud, a cloud-based information system, has been designed and developed by Beihang University to support health education and promotion to the public, restaurants, and food industry; intervention delivery; and project and data management for all programs. This is based on our experience in delivering research projects [24], and on systematic reviews related to nutrition improvement [25] and salt reduction [26] using mHealth technology. The structure of ASCloud is illustrated in Multimedia Appendix 1.

Table 1. Key features of the six programs in Action on Salt China (ASC) at stage 1.

Programs	Purposes	Rationale/design	Coverage at stage 1	Output by the end of stage 1
Program 1: health education and promotion	To improve KAP[a] on salt reduction in the public, restaurants, and food industry; to provide a basis for the other programs	Various types of education materials were developed to improve the KAP targeting major barriers to salt reduction and sources of salt intake using evidence-based key messages	Within study sites of the intervention arms of all 4 RCTs[b] in 6 Chinese provinces[c]	Materials (eg, manuals, fact sheets, leaflets, stickers, public advertisements, short videos, and loudspeaker audio messages) targeted at various populations and settings
Program 2: application-based intervention study (AIS)	To achieve and sustain salt reduction in school children and their families	A cRCT[d] to test the feasibility and effectiveness of an app-based platform (AppSalt) for salt reduction. Goal setting, self-monitoring, and self-reward are the major components [20]	54 primary schools in 3 of the 6 provinces	Finalized AppSalt platform; report on effectiveness of salt reduction as measured by repeated 24-hour urinary sodium excretion; report on feasibility from the perspective of the schools, students, and families
Program 3: home cook-based intervention study (HIS)	To support families, mainly through family cooks, to reduce salt use in home cooking	A cRCT to test the effectiveness and acceptability of a community-based intervention package. Standardized education, salt intake evaluation, individualized recommendations, and reminders are the major components of intervention based on a health belief model [21]	60 communities from the 6 provinces	Intervention package; report on effectiveness of salt reduction measured by 24-hour urinary sodium excretion; report on feasibility
Program 4: restaurant-based intervention study (RIS)	To reduce salt intake when eating out by reducing salt use by restaurant cooks	A cRCT to test the feasibility and effectiveness of a restaurant salt reduction package. Social cognitive theory [22] has been adopted to develop interventions, which include (1) a standardized environment encouraging consumers to order reduced-salt dishes, (2) reminders from waiters, and (3) training cooks to reduce salt use by 10% for all, and greater reduction per consumer requirements	192 restaurants in the 6 provinces	Restaurant intervention package; report on the effectiveness of salt reduction as measured by whole food sodium analysis for each of the restaurants' 5 best-selling dishes; report on feasibility.
Program 5: comprehensive intervention study (CIS)	To explore the experience, barriers, facilitators, and effectiveness of delivering a comprehensive salt reduction intervention	A cRCT at the township/ street level to simulate the scale-up of the intervention clusters and to test its effectiveness. The World Health Organization conceptual framework [23] was adopted to instruct the delivery of all available interventions with close engagement of local government and different sectors.	48 towns /streets in the 6 provinces	Process evaluation report; report of effectiveness of salt reduction as measured by 24-hour urinary sodium excretion.
Program 6: prepackaged food salt reduction	To encourage and support consumers to choose prepackaged foods with lower salt content (P6.1), and to work with the food industry to reduce salt use in prepackaged foods (P6.2)	Besides setting voluntary salt targets[e], consumers are encouraged to choose foods with less salt, and food manufacturers are persuaded to reformulate the products that are high in salt. A health belief model [21] has been used to encourage the consumers to select food with less salt, while convincing food producers that foods high in salt will negatively impact sales.	Participating consumers and food producers	The FoodSwitch[f] app (already downloaded by more than 1 million users); a website designed to raise food manufacturers' awareness of the high salt content of their products and display ranking by salt content in product categories; by late 2019, more than 100 products have already been reformulated to contain less salt

[a]KAP: knowledge, attitude, and practice.

[b]RCT: randomized controlled trial.

[c]The 6 provinces are Heilongjiang, Hebei, Hunan, Jiangxi, Sichuan, and Qinghai, which cover the north, south, central, east, and west part of China.

[d]cRCT: cluster randomized controlled trial.

[e]Setting incremental targets for the salt content of major contributors to salt intake (eg, sauces).

[f]FoodSwitch is a smartphone app that can provide consumers with the nutrition information of a prepackaged food product (including sodium; in China, food products are labeled with sodium rather than salt in which 1 g sodium = 2.5 g salt) and a list of similar food products for making healthier choices, especially with respect to sodium reduction.

Figure 2. The study sites of the Action on Salt China (ASC) cluster of randomized controlled trials.

Program Evaluation

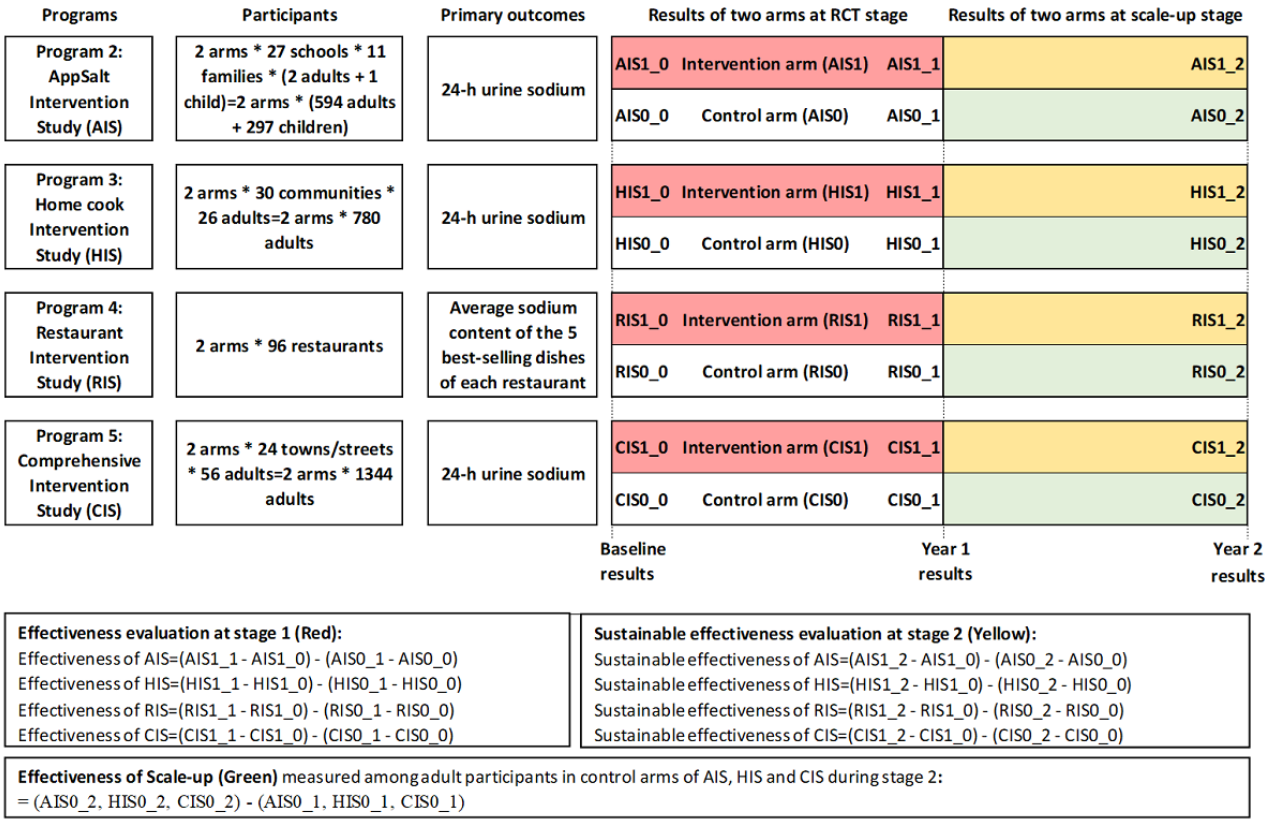
The evaluation method of the effectiveness or impact of the proposed interventions has been specifically designed for each program. In all RCTs, the 1-year effectiveness of the intervention will be assessed by comparing the salt reduction achieved between the intervention and control arms from baseline to the end of stage 1, and the sustained effectiveness over the following year will be assessed by comparing the salt reduction achieved between the intervention and control arms of AIS, HIS, and CIS from baseline to the end of stage 2. Figure 3 shows the design, sample size, and evaluation of primary outcomes for RCTs in Programs 2-5.

The sample sizes indicated in Figure 3 are equal to or slightly larger than the calculated numbers for the RCTs based on target populations (children and adults in AIS, home cooks and family members in HIS, community adults in CIS, and restaurants in RIS), primary outcomes (24-hour sodium excretion in AIS, HIS, and CIS; average sodium content of the 5 best-selling dishes for a restaurant in RIS), standard deviations of primary outcomes (85 mmol/day for 24-hour sodium excretion [19], and 1 g salt/100 g dish [27]), expected minimum salt reduction (25 mmol/day in AIS and CIS; 20 mmol/day for HIS; 0.5 g/100 g dish for RIS [27]), intraclass correlation coefficient (0.05 for AIS, HIS, and CIS [19], and no cluster effect for RIS), drop-out rate (<20%) for individuals/clusters, type I error (5%), and power (>80%), as well as reasonable cluster size (11 students and families in each school in AIS, 13 families in each village in HIS, and 30 adults in each village in CIS).

Secondary outcomes will consist of process evaluation, changes in KAP on salt intake, and economic evaluation. Overall KAP on salt, and overall salt intake levels will also be estimated by pooling the data collected in AIS, HIS, and CIS at baseline, at the end of stage 1, and at the end of stage 2.

Statisticians will be blinded to the intervention assignments during data analysis. An intention-to-treat approach will be adopted for analysis of the primary outcomes. The effect of the intervention on the outcomes will be tested using linear mixed models. For AIS, HIS, and CIS, participants will be nested within family units and families nested within clusters (schools, villages, or towns). Group (intervention, control), time (baseline, end of trial), and time-by-group interaction will be included in the model, with the interaction effect indicating differential change according to group from baseline to the end of the trial. We will adjust for the stratification and potential confounding variables at randomization.

Figure 3. Design for the evaluation of the four cluster randomized controlled trials in Action on Salt China (ASC) Programs 2-5.



Data Collection and Management

Our study will have three major data outputs: (1) data obtained from the evaluation of each RCT's effectiveness (Table 2); (2) data obtained from the process monitoring and evaluation, which will consist of quantitative data automatically generated by the smartphone apps used as intervention tools in the RCTs, as well as qualitative data collected for process evaluation; and (3) monitoring and evaluation data on the coverage and usage of the salt reduction materials and tools, which will be recorded by ASCloud during the scaling-up phase. With the exception of the routine work log and qualitative data collected for process evaluation, most of the data will be collected using specially designed electronic systems, including a mobile device-based electronic data capture system (mEDC) and the ASCloud server, which can capture activities such as log in/log out and access to certain features through different kinds of front ends, Web portals, and mobile apps. The mEDC has an improved process and quality control feature compared with a traditional EDC, and has been validated and widely used in other clinical trials [24,28].

All cleaned and locked datasets for each RCT, together with the study design, questionnaires, code lists, and definitions of databases and variables, will be stored in TGI China, with a unique ID number attached but no personal identifiers, following an established standard operating procedure for data security. To guarantee data security, the mobile app developer (ie, the Information Technology team at Beihang University) will follow the "Mobile Application Information Service Regulation" issued by the Cyberspace Administration of China in 2016 [30]. Although personal data will be accessible to the app developer, disclosure of such information is prohibited.

Table 2. Data collection in the four randomized controlled trials in Action on Salt China (ASC) Programs 2-5.

Questionnaires	Baseline				Year 1 and 2 follow-up
	AIS[a]	HIS[b]	RIS[c,d]	CIS[e]	
Demographics					
Sex	✓	✓	✓	✓	—[f]
Age	✓	✓	✓	✓	—
Education	✓	✓	✓	✓	—
Marriage	✓	✓	✓	✓	—
Income	✓	✓	✓	✓	—
Medical Insurance	✓	✓	✓	✓	—
Knowledge, attitude, and practice (KAP)					
Preference for salt	✓	✓	✓	✓	As baseline
Awareness of salt recommendation	✓	✓	✓	✓	As baseline
Awareness of salt and hypertension	✓	✓	✓	✓	As baseline
Awareness of low-sodium salt	✓	✓	✓	✓	As baseline
Awareness of salt labeling	✓	✓	✓	✓	As baseline
Attitude to low-salt diet	✓	✓	✓	✓	As baseline
Attitude to low-salt behavior	✓	✓	✓	✓	As baseline
Salt use during cooking	✓	✓	✓	✓	As baseline
Frequency of eating out	✓	✓	✓	✓	As baseline
Ordering dishes with reduced salt when eating out	✓	✓	✓	✓	As baseline
Lifestyle					
Smoking	✓	✓	—	✓	As baseline
Drinking	✓	✓	—	✓	As baseline
Physical activity	✓	✓	—	✓	As baseline
Disease history					
Hypertension	✓	✓	—	✓	As baseline
Anti-hypertensive medication use	✓	✓	—	✓	As baseline
Other chronic diseases	✓	✓	—	✓	As baseline
Physical examination					
Height	✓	✓	—	✓	As baseline[g]
Weight	✓	✓	—	✓	As baseline
Waist circumference	✓	✓	—	✓	As baseline
Blood pressure	✓	✓	—	✓	As baseline
Heart rate	✓	✓	—	✓	As baseline
24-hour urinary excretion[h]					
Sodium	✓	✓	—	✓	As baseline
Potassium	✓	✓	—	✓	As baseline
Creatinine	✓	✓	—	✓	As baseline
Albumin	✓	✓	—	✓	As baseline
Calcium	✓	✓	—	✓	As baseline
Restaurants					
Salt-specific environmental factors	—	—	✓	—	As baseline

Questionnaires	Baseline				Year 1 and 2 follow-up
	AIS[a]	HIS[b]	RIS[c,d]	CIS[e]	
Recipe of the 50 best-selling dishes	—	—	✓	—	—
Percentage of consumers who choose lower salt foods	—	—	✓	—	—
Usage of salt and highly salted foods	—	—	✓	—	As baseline
Provision of salt reduction services	—	—	✓	—	As baseline
Whole food laboratory test					
Sodium content of the 5 best-selling dishes	—	—	✓	—	As baseline

[a]AIS: app-based intervention study.

[b]HIS: home cook-based intervention study.

[c]RIS: restaurant-based intervention study.

[d]The primary outcome of RIS is the change of salt use among the study restaurants measured by whole food sodium analysis for the 5 best-selling dishes in each restaurant. Twenty consumers will be invited to take part in a simple survey at baseline and at the end of 2 follow-ups.

[e]CIS: comprehensive intervention study.

[f]Not applicable.

[g]Only the height of children in AIS will be measured during the follow-up visit at year 1 and year 2.

[h]The quality control for 24-hour urine collection refers to the protocol of AIS [29].

Patient and Public Involvement

During development of the overall design of ASC and the specific protocols for each RCT, people from the target populations and those involved in the implementation of the interventions were consulted at least once through meetings, teleconference, and site visits. The consulted individuals included primary school teachers, parents of primary school children, family cooks, consumers of prepackaged food, food producers, restaurant staff, community residents, and policy makers. All ideas on new interventions, opinions on the feasibility of specific interventions, and suggestions to improve their design and implementation were carefully considered. All participants will be informed of the study progress by regular communication via the ASC newsletter, WeChat, and website. Upon completion of the study, we will disseminate the results to all participants and discuss the translation of our study findings to practice.

Results

The duration of ASC is from June 1, 2017 to March 31, 2021, with March 31, 2020 as the split point for stage 1 and stage 2. The preparation of the 4 RCTs and their baseline investigations were completed at the end of March 2019. Protocols of the intervention packages or intervention components that proved to be effective at stage 1 will be made available and scaled up by combining them into an existing national initiative such as Healthy Lifestyle Campaign for All [31] for stage 2. The status of all ASC programs is summarized in Textbox 1.

All trials have been approved by Queen Mary Research Ethics Committee in the United Kingdom (QMERC2018/13 for AIS, QMERC2018/15 for HIS, QMERC2018/16 for CIS, and QMERC2018/14 for RIS) and the Institutional Review Boards of Peking University (IRB00001052-18051 for AIS), Chinese CDC (No. 201801 for HIS), National Center for Chronic and Noncommunicable Disease Control and Prevention, China CDC (No. 201807 for CIS), and National Institute for Nutrition and Health, China CDC (20180314 for RIS). Written informed consent forms have been obtained from all participants according to well-established practices. For children, participant assent and parental written consent have been obtained. All participants are free to discontinue their participation at any time with no explanation required.

By the end of 2019, three steering committee meetings have been convened. Nine presentations have been made in international (four) and national (five) meetings. The Chinese National Health Commission, especially the divisions of disease control, food safety, and health education, is looking forward to adopting the evidence-based salt reduction packages in ASC in 2020, which could help support the Healthy China Initiatives launched in July 2019 [33].

Textbox 1. Status of the programs of Action on Salt China (ASC).

- Program 1: Health education and promotion
  - KAP (knowledge, attitude, and practice) questionnaire: pilot tested and now in use by all RCTs (randomized controlled trial).
  - Education materials: 1 manual, 3 leaflets, 2 public advertisements (15 s and 30 s), 3 short videos targeting 3 major settings for salt intake (home cooking, eating out, and groceries), 8 loudspeaker audio messages for rural village use, and several others; ready for use in RCTs.
- Programs 2-5: app-based intervention study (AIS), home cook-based intervention study (HIS), restaurant-based intervention study (RIS), comprehensive intervention study (CIS)
  - Intervention package development for all 4 RCTs: completed.
  - Participant recruitment: 592 children and 1184 adults recruited for AIS, 1576 home cooks recruited for HIS, 2694 residents recruited for CIS, and 192 restaurants recruited for RIS.
  - Baseline surveys: completed, including 24-hour urine collection in AIS, HIS, and CIS, and whole food sodium analysis for the 5 best-selling dishes of each participating restaurant in RIS.
  - Randomization: Schools/communities/restaurants randomly assigned to intervention and control arms after completion of baseline surveys.
  - Intervention implementation: Ongoing
  - RCT follow-up assessment: completed by the end of January 2020 for AIS, HIS, and CIS, and will be completed by the end of May 2020 for RIS.
- Program 6: Prepackaged food salt reduction
  - FoodSwitch was released June 16, 2019 and has been downloaded by over 10,000 users as of November 2019. It can be found by searching "Shi Xian Zhi" in most app markets and WeChat (a popular communication app in China), and can be used by consumers to choose lower-salt prepackaged foods. Shi Xian Zhi is Chinese pinyin, with Shi indicating food, and Xian Zhi for prophet.
  - ASC's official website now provides a purpose-built page for food producers, based on the data collected through FoodSwitch.
- The information system (ASCloud)
  - ASC's official website [32] was launched in September 2018. To avoid contamination of the RCTs' control arms, it will remain open mainly for internal use and for food producers during stage 1.
  - Several smartphone apps and WeChat applets (ie, very small apps) have also been developed to help consumers choose foods with less salt (FoodSwitch, Shi Xian Zhi in Chinese pinyin), to help families estimate their salt intake (KnowSalt, Jia Ting Yong Yan Ce Liang in Chinese pinyin, used in HIS and CIS), to help deliver a series of health education and activities on salt reduction in schools for the school children and their families (AppSalt, Jian Yan, used in AIS), and to help project implementation and quality control (one app per RCT). Jian Yan is Chinese pinyin, with Jian meaning health and Yan meaning salt.
- Laboratory tests
  - One top-level central laboratory has been contracted for the assay of all urine samples collected in AIS, HIS, and CIS, and the whole food sodium analysis in RIS.

Discussion

Expected Outputs and Potential Impact

ASC is a research unit led by a strong multidisciplinary team with members from international and national institutes engaging in nearly all relevant areas on salt reduction in China. The ASC program is comprehensive, including a salt awareness campaign, legislation support, and RCTs evaluating the strategies dealing with key challenges on salt reduction in China. ASC uses innovative digital health technologies to support the delivery of interventions as well as project and data management. The real-time monitoring and process evaluation can help to increase the fidelity of complex interventions.

As a unit, the six programs of ASC will provide a set of novel approaches to reduce salt intake in China. The expected outputs of ASC include: (1) several evidence-based intervention packages addressing major sources of salt intake; (2) evidence-based salt reduction strategies and experience on policy advocacy along with scaling-up in different regions and populations; and (3) study reports and publications to highlight the gaps, needs, barriers and facilitators, and strategies in salt reduction among different populations. All these major outputs will make significant contributions to the national policies, programs, and initiatives on the prevention and control of noncommunicable diseases as well as the promotion of healthy diets and healthy lifestyle in China. Owing to the strong policy support, multidisciplinary research team, and close partnership with all key national agencies, ASC is most likely to succeed in achieving its intended objectives.

To date, ASC has already made significant progress in achieving its academic impact. As shown in Textbox 1, ASC has attracted the attention of key departments of the Chinese central government. With the support of relevant government agencies, the development of detailed scaling-up plans is underway. If the program is implemented and sustained across China, it will

reduce population salt intake and thereby prevent hundreds of thousands of strokes, heart attacks, and heart failure each year, leading to major cost-savings to individuals, their families, and the health service. Although our study will be carried out in China, the outcomes could potentially be adopted by many other countries. Additionally, our model on salt reduction could possibly be adapted for other dietary and lifestyle changes to prevent CVD and other noncommunicable diseases, which will have major public health implications.

Future Dissemination

The findings of this study will be disseminated widely through conference presentations, peer-reviewed publications, press releases, and social media. In China, the results and effective interventions will be disseminated nationwide through the existing system of health education, disease control, and prevention. Furthermore, the results will be disseminated worldwide through World Action on Salt and Health [34], which is a global nonprofit organization of 600 members from 100 countries with the mission to improve the health of populations by reducing salt intake.

Potential Limitations and Risks

The interventions for each program of ASC are complex. Low compliance to the intervention may lead to negative results for the primary outcomes. In addition, the implementation of collecting 24-hour urine is very demanding. Fidelity to interventions and quality control for effectiveness evaluation are both critical to the success of the programs. With the widespread implementation of Healthy China 2030 initiatives, salt reduction is being promoted by many other programs. Therefore, it might be very difficult to distinguish the contributions of ASC from those of other programs.

Conclusion

The ASC project is progressing smoothly. The intervention packages and components evaluated at stage 1 will provide strong support for salt reduction in China, and could potentially be adopted by many other countries worldwide.

Acknowledgments

The authors would like to thank the staff from the local CDCs, community health service centers, schools, and restaurants, and the participants who shared their opinions on the development of the intervention programs. This research was commissioned by the National Institute for Health Research (NIHR) (NIHR Global Health Research Unit Action on Salt China [ASC] at Queen Mary University of London) using Official Development Assistance (ODA) funding (16/136/77). The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Authors' Contributions

PZ and FH conceived the project and contributed equally to the work. PZ designed the study and drafted the paper together with Xinhua Li, FH, and Yuan Li. CL, JW, JM, BZ, HW, YH Li, J Han, RL, J He, Xian Li, CW, MT, and GM participated in the design of specific programs in this study and reviewed the manuscript. Y Liu led the development of ASCloud, the information system used in this study. Xinhua Li facilitated patient and public involvement and was responsible for setting up the study at each site.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Structure of Action on Salt (ASC) information system (ASCloud) developed by Beihang University. AIS: app-based intervention study; HIS: home cook-based intervention study; RIS: restaurant-based intervention study; CIS: comprehensive intervention study; mEDC: mobile phone app based on an electronic data collection system; PM: project management; MS: management system; NINH: National Institute for Nutrition and Health, China CDC; TGI: The George Institute for Global Health; NCD Division: Division of NCD Control and Community Health, China CDC; NCD Center: National Center for Chronic and Non-communicable Disease Control and Prevention, China CDC; CFSA: China National Center for Food Safety Risk Assessment.
[PNG File , 243 KB - resprot_v9i4e15933_app1.png]

Multimedia Appendix 2
Peer review comments and responses from National Institute for Health Research.
[PDF File (Adobe PDF File), 139 KB - resprot_v9i4e15933_app2.pdf]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1776 KB - resprot_v9i4e15933_app3.pdf]

References

1. World Health Organization. Guideline: Sodium Intake for Adults and Children. Geneva; 2012. URL: https://www.who.int/nutrition/publications/guidelines/sodium_intake/en/ [accessed 2019-12-10]
2. Holbrook JT, Patterson KY, Bodner JE, Douglas LW, Veillon C, Kelsay JL, et al. Sodium and potassium intake and balance in adults consuming self-selected diets. Am J Clin Nutr 1984 Oct;40(4):786-793. [doi: 10.1093/ajcn/40.4.786] [Medline: 6486085]
3. He FJ, MacGregor GA. A comprehensive review on salt and health and current experience of worldwide salt reduction programmes. J Hum Hypertens 2009 Jun;23(6):363-384. [doi: 10.1038/jhh.2008.144] [Medline: 19110538]
4. He FJ, MacGregor GA. Reducing population salt intake worldwide: from evidence to implementation. Prog Cardiovasc Dis 2010;52(5):363-382. [doi: 10.1016/j.pcad.2009.12.006] [Medline: 20226955]
5. Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 Sep 28;394(10204):1145-1158 [FREE Full text] [doi: 10.1016/S0140-6736(19)30427-1] [Medline: 31248666]
6. He FJ, MacGregor GA. Salt reduction lowers cardiovascular risk: meta-analysis of outcome trials. Lancet 2011 Jul 30;378(9789):380-382. [doi: 10.1016/S0140-6736(11)61174-4] [Medline: 21803192]
7. He FJ, Campbell NRC, Ma Y, MacGregor GA, Cogswell ME, Cook NR. Errors in estimating usual sodium intake by the Kawasaki formula alter its relationship with mortality: implications for public health. Int J Epidemiol 2018 Dec 01;47(6):1784-1795 [FREE Full text] [doi: 10.1093/ije/dyy114] [Medline: 30517688]
8. Asaria P, Chisholm D, Mathers C, Ezzati M, Beaglehole R. Chronic disease prevention: health effects and financial costs of strategies to reduce salt intake and control tobacco use. Lancet 2007 Dec 15;370(9604):2044-2053. [doi: 10.1016/S0140-6736(07)61698-5] [Medline: 18063027]
9. Trieu K, Neal B, Hawkes C, Dunford E, Campbell N, Rodriguez-Fernandez R, et al. Salt Reduction Initiatives around the World - A Systematic Review of Progress towards the Global Target. PLoS One 2015;10(7):e0130247 [FREE Full text] [doi: 10.1371/journal.pone.0130247] [Medline: 26201031]
10. He FJ, Pombo-Rodrigues S, Macgregor GA. Salt reduction in England from 2003 to 2011: its relationship to blood pressure, stroke and ischaemic heart disease mortality. BMJ Open 2014 Apr 14;4(4):e004549. [doi: 10.1136/bmjopen-2013-004549] [Medline: 24732242]
11. China GOOSC. China's long-term plan for the prevention and treatment of chronic diseases (2017-2025). 2017 Feb. URL: http://www.nhc.gov.cn/bgt/gwywj2/201702/63b05a3bc7814a3686d5d37f0211f88c.shtml [accessed 2019-12-10]
12. Tan M, He FJ, Wang C, MacGregor GA. Twenty-Four-Hour Urinary Sodium and Potassium Excretion in China: A Systematic Review and Meta-Analysis. J Am Heart Assoc 2019 Jul 16;8(14):e012923. [doi: 10.1161/JAHA.119.012923] [Medline: 31295409]
13. Anderson CAM, Appel LJ, Okuda N, Brown IJ, Chan Q, Zhao L, et al. Dietary sources of sodium in China, Japan, the United Kingdom, and the United States, women and men aged 40 to 59 years: the INTERMAP study. J Am Diet Assoc 2010 May;110(5):736-745 [FREE Full text] [doi: 10.1016/j.jada.2010.02.007] [Medline: 20430135]
14. Zhao F, Zhang P, Zhang L, Niu W, Gao J, Lu L, et al. Consumption and sources of dietary salt in family members in Beijing. Nutrients 2015 Apr 10;7(4):2719-2730 [FREE Full text] [doi: 10.3390/nu7042719] [Medline: 25867952]
15. He FJ, MacGregor GA. Role of salt intake in prevention of cardiovascular disease: controversies and challenges. Nat Rev Cardiol 2018 Jun;15(6):371-377. [doi: 10.1038/s41569-018-0004-1] [Medline: 29713009]
16. CPC Central Committee TSC: 'Healthy China 2030' plan outline. 2016. URL: http://www.gov.cn/gongbao/2016-11/20/content_5133024.htm [accessed 2019-12-12]
17. He FJ, Zhang P, Li Y, MacGregor GA. Action on Salt China. Lancet 2018 Jul 07;392(10141):7-9. [doi: 10.1016/S0140-6736(18)31138-3] [Medline: 30047401]
18. He FJ, Wu Y, Ma J, Feng X, Wang H, Zhang J, et al. A school-based education programme to reduce salt intake in children and their families (School-EduSalt): protocol of a cluster randomised controlled trial. BMJ Open 2013;3(7):e003388 [FREE Full text] [doi: 10.1136/bmjopen-2013-003388] [Medline: 23864214]
19. He FJ, Wu Y, Feng X, Ma J, Ma Y, Wang H, et al. School based education programme to reduce salt intake in children and their families (School-EduSalt): cluster randomised controlled trial. BMJ 2015 Mar 18;350:h770 [FREE Full text] [Medline: 25788018]
20. Payne HE, Moxley VB, MacDonald E. Health Behavior Theory in Physical Activity Game Apps: A Content Analysis. JMIR Serious Games 2015;3(2):e4 [FREE Full text] [doi: 10.2196/games.4187] [Medline: 26168926]
21. Chen J, Liao Y, Li Z, Tian Y, Yang S, He C, et al. Determinants of salt-restriction-spoon using behavior in China: application of the health belief model. PLoS One 2013;8(12):e83262 [FREE Full text] [doi: 10.1371/journal.pone.0083262] [Medline: 24376675]
22. Ahn S, Kwon JS, Kim K, Kim H. Stages of Behavioral Change for Reducing Sodium Intake in Korean Consumers: Comparison of Characteristics Based on Social Cognitive Theory. Nutrients 2017 Jul 27;9(8):808 [FREE Full text] [doi: 10.3390/nu9080808] [Medline: 28749441]
23. Gallani MCBJ, Cornélio ME, Agondi RDF, Rodrigues RCM. Conceptual framework for research and clinical practice concerning cardiovascular health-related behaviors. Rev Lat Am Enfermagem 2013;21 Spec No:207-215 [FREE Full text] [doi: 10.1590/s0104-11692013000700026] [Medline: 23459909]

24. Zhang J, Sun L, Liu Y, Wang H, Sun N, Zhang P. Mobile Device-Based Electronic Data Capture System Used in a Clinical Randomized Controlled Trial: Advantages and Challenges. J Med Internet Res 2017 Mar 08;19(3):e66 [FREE Full text] [doi: 10.2196/jmir.6978] [Medline: 28274907]
25. Li Y, Ding J, Wang Y, Tang C, Zhang P. Nutrition-Related Mobile Apps in the China App Store: Assessment of Functionality and Quality. JMIR Mhealth Uhealth 2019 Jul 30;7(7):e13261 [FREE Full text] [doi: 10.2196/13261] [Medline: 31364606]
26. Ali SH, Luo R, Li Y, Liu X, Tang C, Zhang P. Application of Mobile Health Technologies Aimed at Salt Reduction: Systematic Review. JMIR Mhealth Uhealth 2019 Apr 17;7(4):e13250 [FREE Full text] [doi: 10.2196/13250] [Medline: 30994467]
27. Ma GX, Shive S, Zhang Y, Aquilante J, Tan Y, Zhao M, et al. Knowledge, perceptions, and behaviors related to salt use among Philadelphia Chinese take-out restaurant owners and chefs. Health Promot Pract 2014 Sep;15(5):638-645 [FREE Full text] [doi: 10.1177/1524839914538816] [Medline: 24942751]
28. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
29. He FJ, Zhang P, Luo R, Li Y, Chen F, Zhao Y, et al. An Application-based programme to reinforce and maintain lower salt intake (AppSalt) in schoolchildren and their families in China. BMJ Open 2019 Jul 03;9(7):e027793 [FREE Full text] [doi: 10.1136/bmjopen-2018-027793] [Medline: 31272977]
30. The Cyberspace Administration of China. Mobile Application Information Service Regulation. 2016. URL: http://www.cac.gov.cn/2016-06/28/c_1119123114.htm [accessed 2020-02-17]
31. Li Y, Zhang J, Wang J, Shi X, Liang X. The role of national healthy lifestyle actions in the prevention and control of chronic diseases. Chinese J Prevent Med 2014;48(08):741-743 [FREE Full text] [doi: 10.3760/cma.j.issn.0253-9624.2014.08.020]
32. Action on Salt China. URL: https://www.actionsaltchina.com/ [accessed 2019-12-02]
33. State Council People's Republic of China. Council TS: The opinions of the State Council on the implementation of Health China Action. URL: http://english.gov.cn/policies/latest_releases/2019/07/15/content_281476765851704.htm [accessed 2019-12-09]
34. He FJ, Jenner KH, Macgregor GA. WASH-World Action on Salt and Health. Kidney Int 2010 Oct;78(8):745-753 [FREE Full text] [doi: 10.1038/ki.2010.280] [Medline: 20720531]

Abbreviations

AIS: app-based intervention study
ASC: Action on Salt China
CCHE: Chinese Center for Health Education
CFSA: China National Center for Food Safety Risk Assessment
China CDC: Chinese Center for Disease Control and Prevention
CIS: comprehensive intervention study
CVD: cardiovascular disease
HIS: home cook-based intervention study
KAP: knowledge, attitude, and practice
mEDC: mobile device-based electronic data capture system
NIHR: National Institute for Health Research
RCT: randomized controlled trial
RIS: restaurant-based intervention study
SIPS: scale-up intervention package on salt reduction
TGI: The George Institute
WHO: World Health Organization

Edited by G Eysenbach; submitted 21.08.19; peer-reviewed by S Capewell, P Sever, M Dorsch, M Hanbury; comments to author 28.10.19; revised version received 22.12.19; accepted 13.01.20; published 09.04.20.

Please cite as:
Zhang P, He FJ, Li Y, Li C, Wu J, Ma J, Zhang B, Wang H, Li Y, Han J, Luo R, He J, Li X, Liu Y, Wang C, Tan M, MacGregor GA, Li X
Reducing Salt Intake in China with "Action on Salt China" (ASC): Protocol for Campaigns and Randomized Controlled Trials
JMIR Res Protoc 2020;9(4):e15933
URL: https://www.researchprotocols.org/2020/4/e15933
doi:10.2196/15933
PMID:32271155

Protocol

Effectiveness and Cost-Effectiveness of Receiving a Hearing Dog on Mental Well-Being and Health in People With Hearing Loss: Protocol for a Randomized Controlled Trial

Lucy Stuttard[1], BSc, MSc; Catherine Hewitt[2], PhD; Caroline Fairhurst[2], MSc; Helen Weatherly[3], MSc; Simon Walker[3], MSc; Francesco Longo[3], MSc; Jane Maddison[1], BA; Philip Boyle[1], BA; Bryony Beresford[1], BSc, PhD

[1]Social Policy Research Unit, Department of Social Policy and Social Work, Alcuin B Block, University of York, York, United Kingdom

[2]York Trials Unit, Department of Health Sciences, University of York, York, United Kingdom

[3]Centre for Health Economics, University of York, York, United Kingdom

Corresponding Author:
Lucy Stuttard, BSc, MSc
Social Policy Research Unit
Department of Social Policy and Social Work
Alcuin B Block, University of York
Heslington
York, YO10 5DD
United Kingdom
Phone: 44 01904321965
Email: lucy.stuttard@york.ac.uk

Abstract

Background: People with hearing loss, particularly those who lose their hearing in adulthood, are at an increased risk of social isolation, mental health difficulties, unemployment, loss of independence, risk of accidents, and impaired quality of life. In the United Kingdom, a single third-sector organization provides hearing dogs, a specific type of assistance dog trained to provide sound support to people with hearing loss. These dogs may also deliver numerous psychosocial benefits to recipients. This has not previously been fully investigated.

Objective: The study aims to evaluate the impact of a hearing dog partnership on the lives of individuals with severe or profound hearing loss.

Methods: A 2-arm, randomized controlled trial will be conducted within the United Kingdom with 162 hearing dog applicants, aged 18 years and older. Participants will be randomized 1:1 using a matched-pairs design to receive a hearing dog sooner than usual (intervention arm: arm B) or to receive a hearing dog within the usual timeframe (comparator arm: arm A). In the effectiveness analysis, the primary outcome is a comparison of mental well-being 6 months after participants in arm B have received a hearing dog (arm A have not yet received a hearing dog), measured using the Short Warwick Edinburgh Mental Well-Being Scale. Secondary outcome measures include the Patient Health Questionnaire-9, Generalized Anxiety Disorder-7, and Work and Social Adjustments Scale. An economic evaluation will assess the cost-effectiveness, including health-related quality-adjusted life years using the EuroQol 5 Dimensions and social care–related quality-adjusted life years. Participants will be followed up for up to 2 years. A nested qualitative study will investigate the impacts of having a hearing dog and how these impacts occur.

Results: The study is funded by the National Institute for Health Research's School for Social Care Research. Recruitment commenced in March 2017 and is now complete. A total of 165 participants were randomized. Data collection will continue until January 2020. Results will be published in peer-reviewed journals and at conferences. A summary of the findings will be made available to participants. Ethical approval was received from the University of York's Department of Social Policy and Social Work Research Ethics Committee (reference SPSW/S/17/1).

Conclusions: The findings from this study will provide, for the first time, strong and reliable evidence on the impact of having a hearing dog on people's lives in terms of their quality of life, well-being, and mental health.

Trial Registration: International Standard Randomised Controlled Trial Number Registry ISRCTN36452009; http://www.isrctn.com/ISRCTN36452009

International Registered Report Identifier (IRRID): DERR1-10.2196/15452

(JMIR Res Protoc 2020;9(4):e15452) doi:10.2196/15452

KEYWORDS

randomized controlled trial; hearing loss; qualitative research; economics; assistance dog

Introduction

Background and Rationale

Around 5% of the world's population currently experience disabling hearing loss, and this is estimated to nearly double by 2050 [1]. In the United Kingdom (UK), 1 in 6 adults are affected by hearing loss, and about 1% of the adult population is severely or profoundly deaf. For this latter population, hearing aids offer little benefit [2]. Adults with hearing loss, particularly those who acquired hearing loss in adulthood, are at a risk of adverse outcomes, including social isolation, emotional distress, mental health difficulties, unemployment, dependence, increased risk of accidents, and impaired quality of life [3-10]. Hearing loss is also associated with cognitive decline and an increased risk of dementia [11,12]. People experiencing hearing loss would like access to services, equipment, and assistive techniques that support mental well-being [13], activities of daily living, and the best quality of life [4]. In cases where people are unable to benefit from hearing aids [2], the focus shifts to adaptation of hearing loss and the prevention (or minimization) of adverse outcomes. A hearing dog is one of the support options available [14].

About Hearing Dogs

In the United Kingdom, just one organization, Hearing Dogs for Deaf People [15] trains and provides hearing dogs to UK accredited assistance dog standards [16]. A hearing dog's specialist training means that it can alert a deaf person to a range of everyday sounds (eg, cooker timers and alarm clocks), some of which support communication with others (eg, telephone calls, doorbells, and their name being called) and personal safety (eg, smoke alarms and baby cry monitors). A hearing dog lives with its deaf recipient and, in contrast to a pet dog, is legally permitted (under the Equality Act 2010) to accompany the recipient into service settings (eg, shops, pubs, and aircraft).

Once an application to Hearing Dogs for Deaf People is accepted, a client advisor liaises with the applicant over 3 to 6 months to create a detailed understanding and written *profile* of their lifestyle and hearing dog–relevant needs. This *profile* is then released to the hearing dog training staff who liaise with the client advisor to identify *matches* between an applicant and the pool of hearing dogs coming toward the end of their training (aged approximately 18 months old). It can take between 1 and 3 years for a hearing dog to be provided to an applicant depending on the complexity of their needs. During the period of introducing the hearing dog to the recipient and the dog settling in their home, a series of strategies ensure the hearing dog attaches exclusively to the recipient. Apart from when the recipient is exercising the dog, the hearing dog wears a distinctive *hearing dog jacket* when taken outside the home. Hearing Dogs for Deaf People refers to the arrangement of a hearing dog and recipient as a *hearing dog partnership*. Each recipient is allocated a partnership instructor for ongoing partnership support in their local community. Initially, partnership instructors visit regularly to offer bespoke advice and support, tapering to a minimum annual visit once the partnership is established, although recipients can still contact them for support at any time and attend community activities, events, and workshops. Routine visits ensure the recipient is maintaining the dog's welfare, standard of general behavior, and sound work skills training, thus providing an opportunity for the recipient to request advice and support on these issues or other areas of concern. A hearing dog partnership can last up to 10 years, after which the dog is retired.

Hearing Dogs for Deaf People report that the cost of training a dog and creating and supporting a partnership throughout the dog's working life is around UK £40,000 (US $52,000) [17].

Although Hearing Dogs for Deaf People retain ownership of their hearing dogs and cover the costs described above, recipients are responsible for covering day-to-day costs such as food, routine veterinary care, and insurance.

Existing Evidence

A literature review of evaluations of assistance dogs was conducted to make the case for this research (internal report by Baxter and Beresford [18]). This review identified three studies that had evaluated hearing dog partnerships [19-21] and a further two with mixed samples of individuals with a hearing or mobility impairment [22,23]. These studies indicated that there could be psychological, social, and health benefits of an assistance dog, but the study designs were weak, eg, before and after studies and nonrandomized comparative studies, often with small samples. Only two studies, both mixed samples of individuals with a hearing or mobility impairment, considered the cost implications of a hearing dog. All studies apart from one [19] were conducted in the United States and Canada, where the findings may not be directly transferable to the UK setting. Since the publication of this review, two further studies looking at the impact of hearing dogs on quality of life have been published [24,25]. One of them was a UK study employing a nonrandomized design to compare outcomes for recipients of hearing dogs and mobility dogs to those waiting to receive one [24]. Some improvements in quality of life were reported, but the study had a poor response rate, and the representativeness of the surveyed sample with respect to the population of hearing dog applicants was unclear. The other was a longitudinal study conducted in Sweden of outcomes for 55 dog owners before and after their companion dog was trained to perform as an assistance dog [25]. The majority of these dogs were trained to become mobility (n=30) or diabetes assistance dogs (n=20), with only 3 of them being trained to become hearing dogs. Improvements in quality of life, well-being, and levels of physical activity were reported, but there was no comparator group.

Design

The study will comprise a single-center, superiority randomized controlled trial (RCT) with nested economic and qualitative evaluations. In developing the protocol, the research team worked closely with Hearing Dogs for Deaf People to ensure the design fits with standard processes for creating hearing dog partnerships, the highly variable time it takes to match an applicant with a dog, and Hearing Dogs for Deaf People's commitment to create a partnership within 3 years of application.

Objectives

The overall aim of the study is to evaluate the impact of a hearing dog partnership on the lives of individuals with severe or profound hearing loss.

The study objectives are as follows:

- To determine the impact of a hearing dog partnership on mental well-being, as measured by the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) [26], 6 months post receipt of a hearing dog, compared with applicants who have not yet received their hearing dog.
- To determine the impact of a hearing dog partnership on secondary outcomes of impairment in functioning, anxiety, depression, and health-related and social care–related quality of life 6 months postreceipt of a hearing dog.
- To conduct a nested economic evaluation to investigate the cost-effectiveness of hearing dogs.
- To conduct a nested longitudinal qualitative study to describe the impacts of a hearing dog on recipients' lives, and the mechanisms by which these impacts occur.
- To gather initial data on the long-term outcomes of having a hearing dog.

Methods

Study Setting

The setting is Hearing Dogs for Deaf People, the only organization accredited to provide hearing dogs to UK residents. Research participants are first-time applicants for a hearing dog who may reside anywhere in the United Kingdom.

Inclusion Criterion

The inclusion criterion is first-time applicants for a hearing dog, aged 18 years and older.

Exclusion Criteria

The exclusion criteria include the following:

- Individuals aged 17 years or younger.
- Individuals requiring a dog who can provide sound and vision support.
- Individuals who are replacing a retiring hearing dog.
- Individuals with a learning disability (indicated by the use of a proxy during the application process).
- Individuals still awaiting a hearing dog but whose application is at a stage past the point where randomization could take place. These applicants will have the opportunity to complete the research materials and form a part of an exploratory longitudinal cohort study.

Intervention

The intervention includes the receipt of a hearing dog specifically matched and trained to support the needs of the applicant. The comparator includes not having a hearing dog. Practices regarding cessation of a hearing dog partnership should align with the standard Hearing Dogs for Deaf People protocols. Outside of the intervention, participants should receive care as usual.

Recruitment

Recruitment to the trial took place between March 2017 and March 2018. Hearing Dogs for Deaf People will screen all applications during this period. Typically, the charity receives over 200 applications annually. For those fulfilling the study eligibility criteria, Hearing Dogs for Deaf People will post a study recruitment pack (including a participant information sheet, consent form, contact preferences form, study questionnaire, and reply-paid envelope). Applicants wishing to participate can choose to complete the study materials on paper or on the Web in either English or British Sign Language (BSL) via Qualtrics, a Web-based survey platform. These will be returned directly to the research team. In the case of nonresponse, Hearing Dogs for Deaf People will send up to two reminders (text message and email or post). Where possible, reasons for decline will be obtained. Applicants not recruited to the study will follow standard Hearing Dogs for Deaf People procedures and timelines.

Randomization

Randomization will be conducted centrally by the York Trials Unit (YTU), using an allocation schedule generated in Stata v15. During their hearing dog assessment, study participants will be categorized by a senior practitioner within Hearing Dogs for Deaf People, with regard to their presenting needs, as follows:

- None: no remarkable or particular needs.
- Personal: predominantly personal needs, this might include particular health concerns or mobility.
- Environmental: predominantly environmental needs, this might include an inner city location or the presence of cats in the home.
- Both: significant personal and environmental needs.

When the profile of 2 individuals with the same categorization of need is completed, they will form a pair and be randomized together, one to each group using block randomization with a block size of 2:

- Arm A: matching with a hearing dog occurs within usual timelines or
- Arm B: matching accelerated, receive a hearing dog at least six months earlier than those in arm A.

The allocation sequence will be generated by the trial statistician (CF) who has no involvement in the recruitment of participants. As pairwise randomization is being employed, ie, the randomization of 2 participants at a time, allocation is concealed.

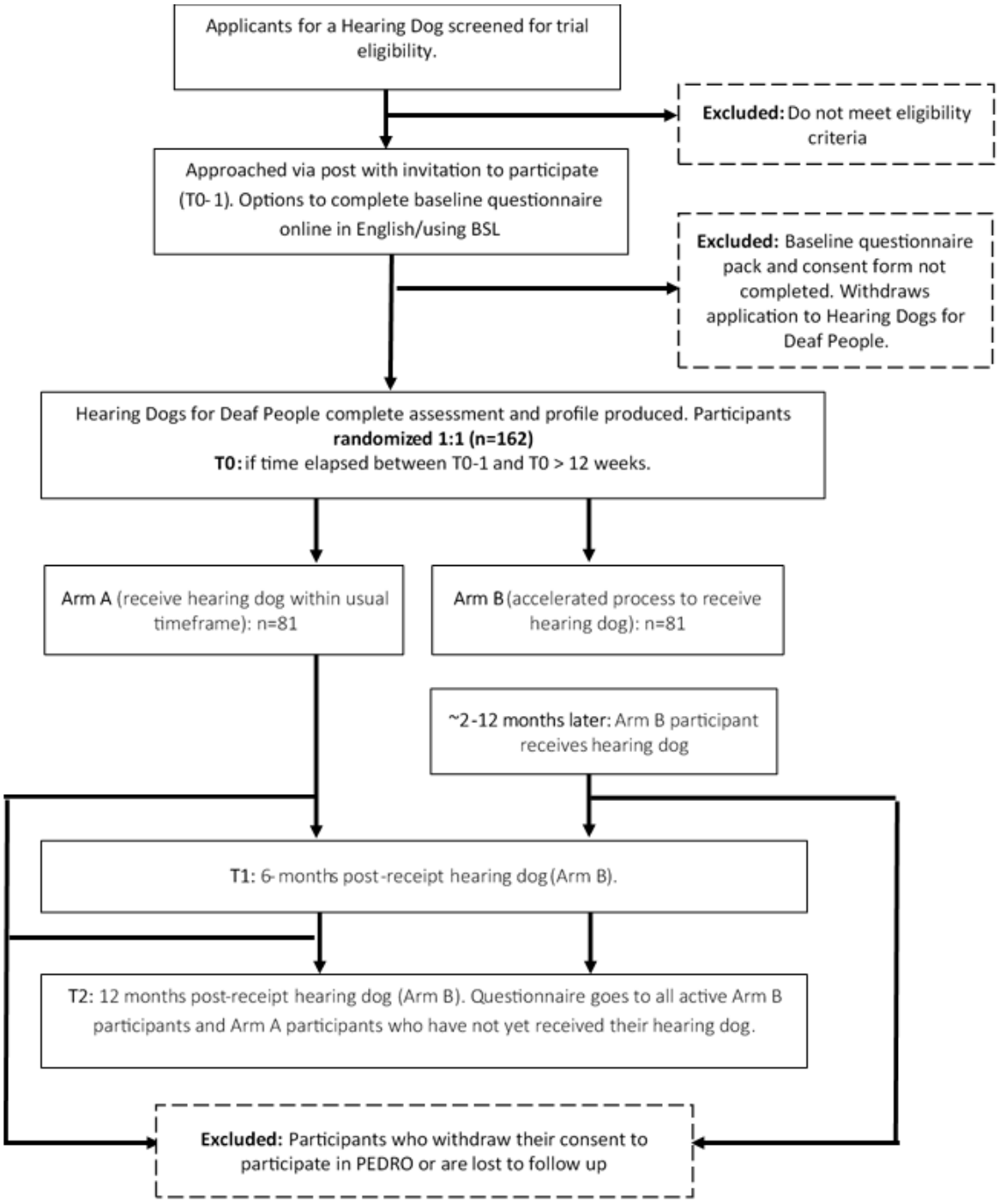
Participants will be blinded to their group allocation. Study team members who are actively involved in the administration of the trial will not be blinded.

Follow-Up Data Collection

Follow-up questionnaires will be administered by the research team via post or email according to participants' preferences. Postal, email, and text reminders and an incentive (£20 per data collection time point) will support retention to the study.

Figure 1 presents the flow of study participants through the trial. T1 (6 months after the arm B participant has received their hearing dog) is the primary outcome time point. Arm B participants will also be followed up at 12 (T2), 18 (T3), and 24 (T4) months post receipt of a hearing dog. Arm A participants will only be followed up at T2 if they have not yet received their dog. Participants who withdraw their application or return their dog will continue to be contacted according to this schedule. In the case of arm B participants withdrawing, a dummy date of partnership will be created. Given the personalized nature of the intervention, some arm A participants may receive their dog before their arm B partner.

Figure 1. Flow of trial participants through partnerships between deaf people and hearing dogs.



Outcome Measures

The selection of outcome measures was informed by the existing literature and in consultation with Hearing Dogs for Deaf People and hearing dog recipients. All standardized measures are available in BSL, and these versions have satisfactory psychometric properties [27-29].

Primary Outcome for the Effectiveness Analysis

The SWEMWBS [26] comprises seven positively worded items related to psychological functioning with five response categories (*none of the time* to *all of the time*). Respondents indicate the response that best describes their experience over the last 2 weeks.

Primary Outcome for the Cost-Effectiveness Analysis

The EuroQol 5 Dimensions (EQ-5D-5L) is a measure of health-related quality of life capturing five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [30]. Each domain is captured on a 5-point scale, with respondents reporting how they feel *today*.

Secondary Outcomes

The Work and Social Adjustments Scale (WSAS) measures impairment in functioning and comprises five items, one related to work and the remainder to social functioning [31]. Respondents rate their impairment on a 9-point scale (*not at all* to *very severely*).

Two measures of mental health will be used: the Patient Health Questionnaire (PHQ-9) and the Generalized Anxiety Disorder Assessment (GAD-7) [32,33]. Respondents rate how often they have experienced numerous mental health problems over the previous 2 weeks. Each item provides a 4-point response scale (*not at all* to *nearly every day*). Clinical cutoffs for BSL versions have been established [34].

Descriptive outcomes include the status of the hearing dog partnership (intact vs ceased); use of statutory, third sector, and private health and social care services in the past 4 weeks; and employment status.

The Hearing Dog Questionnaire [19] will be used to capture reports of problems associated with hearing impairment that a hearing dog is meant to alleviate: awareness of sounds (eg, smoke alarm and doorbell), concerns about security (in and outside the home), dependency on others, and being misunderstood when out in public. It comprises 11 items, and a 5-point scale captures the frequency at which these difficulties are experienced.

Study participants who receive a hearing dog during the trial will also be asked about the benefits and challenges of their hearing dog partnership using fixed response (eg, "How challenging did you find adjusting to having a hearing dog?" with four response options from *not at all challenging* to *very challenging*) and free-text response formats (eg, Please tell us briefly how your life has improved since having a hearing dog).

Data Management

Data will be entered into SPSS 23 by someone who is independent of the data analysis. A random sample of 10% will be double entered to assess data quality. An error rate exceeding 5% will require investigation including an examination of the type of errors, eg, random vs systematic and potentially a double entry of the whole data set. Missing data from a scale will be managed using guidance from scale developers. For the SWEMWBS and WSAS, up to one missing item and, for GAD-7 and PHQ-9, up to two missing items will be replaced with the scale mean.

Statistical Methods

Sample Size

There are limited published data on which to base a sample size for this trial. Therefore, we have taken a pragmatic approach and calculated a sample size that should be achievable within the study timescale. We will aim to approach at least 200 applicants, of which we expect approximately 180 to be eligible and of which 90.0% (162/180) will opt to join the study. Allowing for 20% attrition at T1, this will result in a trial sample at T1 of 128. This pragmatic sample size will provide 80% power at 5% significance to detect an effect size of 0.5 of a standard deviation on our primary outcome measure, the SWEMWBS. This is a moderate-to-large effect size.

Effectiveness Analysis

Trial analyses on the effectiveness of hearing dogs will be conducted on an intention-to-treat basis, including all participants in the arm to which they were randomized. Analyses will be conducted in SPSS 23 or later, using 2-sided statistical tests at the 5% significance level. The flow of participants through each stage of the trial will be presented in a Consolidated Standards of Reporting Trials diagram. The primary analysis will estimate the difference in the SWEMWBS scores at T1 between arms A and B using linear regression, adjusting for the baseline measure and the individual's presenting needs (identified before randomization). The difference between arms in the SWEMWBS score at T1 and corresponding 95% CI will be presented. Model assumptions will be checked, and if they are in doubt, data transformations, alternative distributions, or nonparametric alternatives will be considered. The secondary outcomes (WSAS, PHQ-9, and GAD-7) will be analyzed in the same way.

Sensitivity analyses will investigate the robustness of the findings given any nonadherence to the protocol. Application withdrawals, placement cessations, and adverse events will be summarized for each group.

An exploratory analysis of outcomes at T2 will analyze the data using the same model as for the T1 with the caveat that there will be systematic missing data as T2 data are not collected from arm A participants who have received a hearing dog by this point. Further exploratory analyses will examine the longer-term outcomes collected for study participants where data collection at 18 (T3) and 24 (T4) months post receipt of a hearing dog is achieved. For these analyses, time will be added as a variable to the regression model.

To date, the properties of the Hearing Dog Questionnaire have not been tested to determine whether the items perform as a scale. We will use classical test theory [35] to determine whether

it is appropriate to calculate a total or subscale scores. If this is not possible, we will present responses to the individual items.

The Statistical Analytical Plan will be uploaded to the trial registry.

Economic Evaluation

An economic evaluation will be conducted to determine whether hearing dogs are value for money. It will be undertaken from multiple perspectives to inform value for money considerations for different potential decision makers. Perspectives will include (1) a voluntary sector perspective (including costs to Hearing Dogs for Deaf People and of volunteers); (2) a social care perspective (considering that costs of hearing dogs could fall on social care budgets); (3) a health care perspective (considering the costs of hearing dogs could fall on health care budgets); (4) a public sector perspective (considering impact on both social and health care budgets); and (5) a broader perspective considering costs to the voluntary sector, social care, health care, and costs incurred by the recipients of hearing dogs. The choice of outcome will be perspective dependent, choosing the most relevant outcome(s) for each decision maker, with social care quality-adjusted life years (SC-QALYs) and health quality-adjusted life years (H-QALYs) being key outcomes to consider. To determine cost-effectiveness, incremental costs and units of effect will be compared with relevant cost-effectiveness thresholds [36-38] using incremental cost-effective ratio decision rules and net benefit decision rules where appropriate [39].

Data on costs and resource use will be collected from Hearing Dogs for Deaf People (via structured interview and documentary analysis) and study participants at each data collection point (using a previously developed service and resource use questionnaire that will be updated specifically for this project). Costs will be calculated by applying unit costs to resource use. National unit costs (eg, the Personal Social Service Research Unit's unit costs of health and social care) [40] will be used, where available, to aid the generalizability of findings. The cost of the service will be calculated and reported as an average cost of a hearing dog user.

For outcomes, data on health-related quality of life will be collected using the EQ-5D-5L questionnaire at multiple time points. SC- and H-QALYs will be estimated using these data. H-QALYs will be calculated using the EQ-5D-5L score [41] and the area under the curve (AUC) method [39]. The SC-QALYs will be derived by converting the EQ-5D-5L's answers into EQ-5D-3L's [42], as per the current recommendation by the National Institute for Health and Care Excellence [43], and then by applying the exchange rate proposed by Stevens et al [44] to obtain A Severity Characterization of Trauma (ASCOT) score to be used for the AUC method.

A regression analysis will be undertaken to account for any baseline differences in study participants between the trial arms using appropriate techniques to account for non-normality of outcomes and costs data [45]. Decision uncertainty will be addressed using probabilistic sensitivity analysis, and deterministic sensitivity analysis will be used to examine the impact on results of varying relevant parameters and assumptions.

Nested Qualitative Study

The qualitative study seeks to understand the *active ingredients* of a hearing dog partnership, outcomes of partnerships (positive and negative, expected and unanticipated), the processes by which changes in outcomes (or not) are perceived to occur, and views on the process by which partnerships are created and supported.

It will include a longitudinal study of 15 study participants who have received their hearing dog. This number provides pragmatic balance, allowing both for (1) recruitment to cover the range of variables in the purposive sampling frame (see below) and (2) the capacity and resource demands of the longitudinal approach. A purposive sampling frame will be used (based on, eg, age, gender, age at onset of hearing impairment, presenting health and social needs, family composition, and previous experience of a dog as a pet) to ensure a range of factors and circumstances are represented. Face-to-face interviews will take place approximately 4 and 10 months post receipt of a hearing dog. Participants will be able to participate in English or BSL. Depending on the language used, we will seek permission to audio or video record the interview. For interviews conducted in BSL, a detailed summary of the interview will be produced by the interviewer, an approach we have used successfully in the past [46]. Interviews will explore experiences of the introduction of the hearing dog into the household, perceived impacts on self and wider family, views on factors that have supported or hindered the development of the partnership and its potential impact, and views and experiences of the application and matching process. During their second interview, recipients with a resident partner will be asked for permission to approach them regarding participation in a study interview. Thus, up to 15 partners (permanently living in the same household) will be recruited. Telephone or face-to-face interviews will be used to explore experiences of the introduction of the hearing dog into the household, perceived impacts (positive and negative) on self and hearing dog recipients and on the wider family, views on factors that have supported or hindered the development of the partnership and its potential impact, and experiences of the application and matching process.

A cross-sectional qualitative study will explore the views and experiences of Hearing Dogs for Deaf People staff regarding factors that hinder or facilitate positive outcomes of the hearing dog partnership. Focus groups (n=7; 5-10 participants per group) will be used to gather views and experiences of the three groups of staff most involved with hearing dog applicants and recipients: Client advisors (who support an applicant through to their *match* with a hearing dog, n=1 focus group); dog trainers (involved in matching and handover of the hearing dog to a new recipient, n=2 focus groups), and partnership instructors (responsible for ongoing recipient support, n=4 focus groups). The focus group with client advisors will explore preparation of applicants for a hearing dog, applicants' expectations and concerns, and factors affecting efficiency and quality of the matching process. Focus groups with trainers will explore dog and person characteristics, early indicators of a *successful match*

and the *active ingredients* of the intervention. Focus groups with partnership instructors will explore factors that hinder or facilitate positive outcomes of hearing dog partnerships, the maintenance of the partnership, what constitute the *active ingredients* of the intervention, and experiences of interfacing or collaborating with statutory services. The number of focus groups per staff group reflects differences in size of each workforce. Focus groups for trainers and partnership instructors will take place in both of Hearing Dogs for Deaf People's training bases (South and North England). Overall, the number and location of groups will allow the research team to collect data of sufficient breadth and depth.

The data will be analyzed in two ways. First, for the interviews with recipients, we will create a narrative account that records their experiences of a hearing dog partnership and traces changes in outcomes and life situations of interviewees' lives, which are in some way attributed to the partnership, and the factors that were perceived to precipitate, support, or hinder those changes. These narratives will then be collectively interrogated using thematic analytical techniques to identify and describe the *active ingredients* of a hearing dog partnership and the factors that support and hinder the effectiveness of that partnership. We have used this approach to analyzing longitudinal interviews in the past and found it a very effective and efficient tool [47].

Second, thematic analysis [48] of interviews with recipients, their partners, and Hearing Dogs for Deaf People staff will be used to identify and describe views and experiences regarding the process of creating and supporting partnerships, impacts and consequences, factors perceived to facilitate or hinder positive outcomes, and views regarding the *active ingredients* of a hearing dog partnership. An additional theme—perceived impacts on the self—will be explored for the partner interviews. For the analysis, we will use the Framework approach [49], which facilitates systematic data management and an audit trail of the analytical process.

Patient and Public Involvement

A user advisory panel (UAP) comprising 10 individuals with a hearing dog has been formed. The panel will meet virtually using an online forum and group email. We will seek the views of the UAP on all aspects of the project, particularly the design and content of study information and consent materials, analysis of qualitative data, interpretation and synthesis of findings, and the dissemination strategy. Discussions with the UAP will be shared with the study steering committee. Hearing Dogs for Deaf People will also be consulted about dissemination and impact pathways, and we will advise them regarding their own dissemination of study findings.

Independent Oversight of the Study

A study steering committee has been appointed comprising academics, researchers, and hearing dog recipients. The steering group will meet on four occasions over the course of the study.

Ethics and Dissemination

The study protocol (v3 12/02/2019) included the original application and subsequent amendments (as required) received a favorable ethical opinion from the University of York Department of Social Policy and Social Work Research Ethics Committee (SPSW/S/17/1).

The study has been designed so that no participant will wait longer to receive a hearing dog than is usually expected. The nature of the two arms and the process of randomization will be made clear in study information materials as will the fact that participation is voluntary and choosing not to participate will not affect their application to Hearing Dogs for Deaf People or the service they receive. Study information materials will be developed in consultation with the user advisory panel.

Consent

All participants will be required to provide written, informed consent on joining the study. Separate consent will be obtained for the nested qualitative study. Participants will be informed they can withdraw from the study at any time without this influencing their application.

Confidentiality

All study-related information will be stored securely at the University of York. Participant information will be stored in locked filing cabinets in areas with limited access. All outcomes data will be anonymized and given a coded ID. Records that contain names or other personal identifiers such as address and consent forms will be stored separately from outcomes data. Electronic data will be saved on a secure university filespace, access restricted to members of the research team that have a role analysis. Following the study, all data will be archived and stored in accordance with the University of York guidelines.

Protocol Amendments

All amendments to the protocol will be approved by the research ethics committee.

Dissemination Policy

A summary of the findings will be published by the National Institute of Health Research's (NIHR) School for Social Care Research. Study findings will also be reported in open access, peer-reviewed journals and at relevant conferences. Participants will be sent a summary of the findings.

Monitoring

Data Monitoring

Given the nature of the trial, the trial does not have a data monitoring committee.

Adverse Events

The study will record and report any details of serious adverse events.

Auditing

This trial does not have an audit procedure in place.

Access to Data

The trial manager (LS) will oversee access to data. Cleaned data sets will be shared with the PI (BB), supervising statisticians (CF and CH), and economic team (HW, SW, and FL). Data dispersed to the research team will be blinded of any identifying participant information.

Ancillary and Posttrial Care

Not applicable for this study.

Results

The study is funded by an NIHR School for Social Care grant, and recruitment commenced in March 2017. Recruitment is now complete, and 165 participants were randomized. Data collection is ongoing.

Discussion

This is the first time an RCT design has been used to evaluate the impacts of the hearing dog partnership. The findings from this study will provide, for the first time, strong and reliable evidence on the impact of having a hearing dog on people's lives in terms of their quality of life, well-being, and mental health.

In addition, we have shown that it is possible to do research that collects robust data on the impacts of assistance dogs on people's lives. We think this study will encourage further research in this area, including for other types of assistance dogs.

The study findings will be of relevance to people with hearing impairment, Hearing Dogs for Deaf People and its supporters, and statutory services responsible for the care and support of people with hearing impairments (eg, audiology services and social work or social care services). Findings will support evidence-informed policy making, service development, and practice. Study findings will also be relevant to assistance dog organizations, in the United Kingdom and elsewhere.

Acknowledgments

The authors would like to thank the members of the study steering committee and advisory panel, the study participants for their ongoing participation in the study, and the project administrator. The authors would also like to thank Hearing Dogs for Deaf People without whose cooperation and commitment the study would not have been possible. This work is supported by the NIHR's School for Social Care Research grant number C088/CM/UYYB-P111. The University of York is the study sponsor and has legal responsibility for the initiation and management of the study: sponsor representative Mr Michael Barber, Research and Enterprise Directorate, Innovation Office, York Science Park, Heslington, York, YO10 5DG, United Kingdom. The study sponsor and funder have no role in the design of this study and will not have any role during its execution, analysis, interpretation of the data, or decision to publish results. The views expressed are those of the author(s) and not necessarily those of Hearing Dogs for Deaf People or the NIHR.

Authors' Contributions

BB is the principal investigator at Social Policy Research Unit (SPRU), University of York; CH is responsible for the trial design at YTU, University of York; CF is the statistician at YTU; LS is the trial manager, SPRU; HW is a health economist at the Centre for Health Economics (CHE), University of York; SW is a health economist at CHE; FL, CHE; JM is the qualitative researcher (SPRU); and PB is the study support officer. LS led the preparation of the manuscript, and BB, CF, JM, and HW commented on drafts. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer-review reports from the National Institute for Health Research - Part 1.
[PDF File (Adobe PDF File), 455 KB - resprot_v9i4e15452_app1.pdf]

Multimedia Appendix 2
Peer-review reports from the National Institute for Health Research - Part 2.
[PDF File (Adobe PDF File), 385 KB - resprot_v9i4e15452_app2.pdf]

Multimedia Appendix 3
Peer-review reports from the National Institute for Health Research - Part 3.
[PDF File (Adobe PDF File), 112 KB - resprot_v9i4e15452_app3.pdf]

References

1. World Health Organization. 2019 Mar 20. Deafness and Hearing Loss URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss [accessed 2019-11-21]
2. Turton L, Smith P. Prevalence & characteristics of severe and profound hearing loss in adults in a UK National Health Service clinic. Int J Audiol 2013 Feb;52(2):92-97. [doi: 10.3109/14992027.2012.735376] [Medline: 23205712]

3. Carlsson PI, Hjaldahl J, Magnuson A, Ternevall E, Edén M, Skagerstrand A, et al. Severe to profound hearing impairment: quality of life, psychosocial consequences and audiological rehabilitation. Disabil Rehabil 2015;37(20):1849-1856. [doi: 10.3109/09638288.2014.982833] [Medline: 25391816]
4. Department of Health. NHS England. 2015. Action Plan on Hearing Loss URL: https://www.england.nhs.uk/wp-content/uploads/2015/03/act-plan-hearing-loss-upd.pdf [accessed 2019-11-21]
5. Lustig TA, Domnitz S, Cilio CM, Wunderlich G, Pope AM. Forum on aging, disability, and independence. In: Aging and Disability: Beyond Stereotypes to Inclusion: Proceedings of a Workshop. Washington, DC: The National Academies Press; 2014.
6. Davis A, Smith P, Ferguson M, Stephens D, Gianopoulos I. Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models. Health Technol Assess 2007 Oct;11(42):1-294 [FREE Full text] [doi: 10.3310/hta11420] [Medline: 17927921]
7. Hallam R, Ashton P, Sherbourne K, Gailey L. Acquired profound hearing loss: mental health and other characteristics of a large sample. Int J Audiol 2006 Dec;45(12):715-723. [doi: 10.1080/14992020600957335] [Medline: 17132560]
8. Ringdahl A, Grimby A. Severe-profound hearing impairment and health-related quality of life among post-lingual deafened Swedish adults. Scand Audiol 2000;29(4):266-275. [doi: 10.1080/010503900750022907] [Medline: 11195947]
9. Thomas AJ. Acquired Hearing Loss: Psychological and Psychosocial Implications. London: Academic Press; 1985.
10. Shield B. Hear-it. London: London South Bank University; 2006 Oct. Evaluation of the Social and Economic Costs of Hearing Impairment: A Report for Hear-It URL: https://www.hear-it.org/sites/default/files/multimedia/documents/Hear_It_Report_October_2006.pdf [accessed 2019-11-21]
11. Lin FR, Yaffe K, Xia J, Xue Q, Harris TB, Purchase-Helzner E, Health ABC Study Group. Hearing loss and cognitive decline in older adults. JAMA Intern Med 2013 Feb 25;173(4):293-299 [FREE Full text] [doi: 10.1001/jamainternmed.2013.1868] [Medline: 23337978]
12. Action on Hearing Loss. London: UCL Deafness Cognition and Language Research Centre; 2013. Joining Up: Why People With Hearing Loss or Deafness Would Benefit From an Integrated Response to Long-term Conditions URL: https://www.actiononhearingloss.org.uk/about-us/our-research-and-evidence/research-reports/joining-up-report/ [accessed 2019-11-21]
13. Department of Health. London; 2015. Mental Health and Deafness: Towards Equity in Access URL: http://www.deafinfo.org.uk/policy/Towards%20Equity%20and%20Access.pdf [accessed 2019-11-21]
14. Hearing Dogs for Deaf People. What Are the Impacts of Deafness? URL: https://www.hearingdogs.org.uk/helping-deaf-people/ [accessed 2019-01-22]
15. Hearing Dogs for Deaf People. URL: https://www.hearingdogs.org.uk/ [accessed 2020-01-22]
16. Assistance Dogs UK. URL: http://www.assistancedogs.org.uk/ [accessed 2020-01-22]
17. Hearing Dogs for Deaf People. We Train Hearing Dogs Who Transform Deaf People's Lives URL: https://www.hearingdogs.org.uk/about/ [accessed 2020-01-22]
18. Baxter K, Beresford B. University of York. 2016 Oct. A Review of Methods of Evaluation and Outcome Measurement of a Complex Intervention in Social Care: the Case of Assistance Dogs URL: https://www.york.ac.uk/media/spru/projectfiles/SSCR%202672%20(Assistance%20Dogs%20Review%20-%20internal%20report.pdf [accessed 2020-02-03]
19. Guest CM, Collis GM, McNicholas J. Hearing dogs: a longitudinal study of social and psychological effects on deaf and hard-of-hearing recipients. J Deaf Stud Deaf Educ 2006;11(2):252-261. [doi: 10.1093/deafed/enj028] [Medline: 16452611]
20. Hart LA, Zasloff RL, Benfatto AM. The pleasures and problems of hearing dog ownership. Psychol Rep 1995 Dec;77(3 Pt 1):969-970. [doi: 10.2466/pr0.1995.77.3.969] [Medline: 8559940]
21. Hart LA, Zasloff RL, Benfatto AM. The socializing role of hearing dogs. Appl Anim Behav Sci 1996;47(1-2):7-15. [doi: 10.1016/0168-1591(95)01006-8]
22. Rintala DH, Matamoros R, Seitz LL. Effects of assistance dogs on persons with mobility or hearing impairments: a pilot study. J Rehabil Res Dev 2008;45(4):489-503 [FREE Full text] [doi: 10.1682/jrrd.2007.06.0094] [Medline: 18712636]
23. Valentine DP, Kiddoo M, LaFleur B. Psychosocial implications of service dog ownership for people who have mobility or hearing impairments. Soc Work Health Care 1993;19(1):109-125. [doi: 10.1300/j010v19n01_07]
24. Hall SS, MacMichael J, Turner A, Mills DS. A survey of the impact of owning a service dog on quality of life for individuals with physical and hearing disability: a pilot study. Health Qual Life Outcomes 2017 Mar 29;15(1):59 [FREE Full text] [doi: 10.1186/s12955-017-0640-x] [Medline: 28356121]
25. Lundqvist M, Levin L, Roback K, Alwin J. The impact of service and hearing dogs on health-related quality of life and activity level: a Swedish longitudinal intervention study. BMC Health Serv Res 2018 Jun 27;18(1):497 [FREE Full text] [doi: 10.1186/s12913-018-3014-0] [Medline: 29945630]
26. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. Health Qual Life Outcomes 2009 Feb 19;7:15 [FREE Full text] [doi: 10.1186/1477-7525-7-15] [Medline: 19228398]
27. Rogers KD, Young A, Lovell K, Campbell M, Scott PR, Kendal S. The British Sign Language versions of the Patient Health Questionnaire, the Generalized Anxiety Disorder 7-item Scale, and the Work and Social Adjustment Scale. J Deaf Stud Deaf Educ 2013 Jan;18(1):110-122 [FREE Full text] [doi: 10.1093/deafed/ens040] [Medline: 23197315]

28. Rogers KD, Pilling M, Davies L, Belk R, Nassimi-Green C, Young A. Translation, validity and reliability of the British Sign Language (BSL) version of the EQ-5D-5L. Qual Life Res 2016 Jul;25(7):1825-1834 [FREE Full text] [doi: 10.1007/s11136-016-1235-4] [Medline: 26887955]
29. Rogers KD, Dodds C, Campbell M, Young A. The validation of the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) with deaf British sign language users in the UK. Health Qual Life Outcomes 2018 Jul 24;16(1):145 [FREE Full text] [doi: 10.1186/s12955-018-0976-x] [Medline: 30041627]
30. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
31. Mundt JC, Marks IM, Shear MK, Greist JM. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br J Psychiatry 2002 May;180:461-464. [doi: 10.1192/bjp.180.5.461] [Medline: 11983645]
32. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
33. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
34. Belk RA, Pilling M, Rogers KD, Lovell K, Young A. The theoretical and practical determination of clinical cut-offs for the British Sign Language versions of PHQ-9 and GAD-7. BMC Psychiatry 2016 Nov 3;16(1):372 [FREE Full text] [doi: 10.1186/s12888-016-1078-0] [Medline: 27809821]
35. DeVellis RF. Classical test theory. Med Care 2006 Nov;44(11 Suppl 3):S50-S59. [doi: 10.1097/01.mlr.0000245426.10853.30] [Medline: 17060836]
36. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health Technol Assess 2015 Feb;19(14):1-503, v-vi [FREE Full text] [doi: 10.3310/hta19140] [Medline: 25692211]
37. Forder J, Malley J, Towers A, Netten A. Using cost-effectiveness estimates from survey data to guide commissioning: an application to home care. Health Econ 2014 Aug;23(8):979-992. [doi: 10.1002/hec.2973] [Medline: 24038337]
38. Thokala P, Ochalek J, Leech AA, Tong T. Cost-effectiveness thresholds: the past, the present and the future. Pharmacoeconomics 2018 May;36(5):509-522. [doi: 10.1007/s40273-017-0606-1] [Medline: 29427072]
39. Drummond MF, Sculpher MJ, O'Brien BJ, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Fourth Edition. Oxford: Oxford University Press; 2015.
40. Curtis LA, Burns A. Unit Costs of Health and Social Care 2018. Kent: University of Kent; 2018.
41. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
42. van Hout B, Janssen M, Feng Y, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012;15(5):708-715 [FREE Full text] [doi: 10.1016/j.jval.2012.02.008] [Medline: 22867780]
43. National Institute for Clinical Excellence. Position Statement on Use of the EQ-5D-5L Value Set for England (updated October 2019) URL: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l [accessed 2020-01-22]
44. Stevens K, Brazier J, Rowen D. Estimating an exchange rate between the EQ-5D-3L and ASCOT. Eur J Health Econ 2018 Jun;19(5):653-661 [FREE Full text] [doi: 10.1007/s10198-017-0910-x] [Medline: 28623464]
45. Barber J, Thompson S. Multiple regression of cost data: use of generalised linear models. J Health Serv Res Policy 2004 Oct;9(4):197-204. [doi: 10.1258/1355819042250249] [Medline: 15509405]
46. Greco V, Beresford B, Sutherland H. Deaf children and young people's experiences of using specialist mental health services. Child Soc 2009;23(6):455-469. [doi: 10.1111/j.1099-0860.2008.00176.x]
47. Maddison J, Beresford B. The development of satisfaction with service-related choices for disabled young people with degenerative conditions: evidence from parents' accounts. Health Soc Care Community 2012 Jul;20(4):388-399. [doi: 10.1111/j.1365-2524.2011.01042.x] [Medline: 22360567]
48. Miles MB, Huberman MA. Qualitative Data Analysis: An Expanded Sourcebook. London: Sage Publications; 1994.
49. Ritchie J, Lewis J. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: Sage Publications; 2003.

Abbreviations

ASCOT: A Severity Characterization of Trauma
AUC: area under the curve
BSL: British Sign Language
CHE: Centre for Health Economics
EQ-5D-5L: EuroQol 5 Dimensions
GAD: Generalized Anxiety Disorder

H-QALY: health quality-adjusted life year
NIHR: National Institute of Health Research
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial
SPRU: Social Policy Research Unit
SC-QALY: social care quality-adjusted life year
SWEMWBS: Short Warwick-Edinburgh Mental Well-Being Scale
UAP: user advisory panel
WSAS: Work and Social Adjustments Scale
YTU: York Trials Unit

Edited by G Eysenbach; submitted 18.07.19; peer-reviewed by M Pilling, M Hallewell, Y Chu; comments to author 21.10.19; revised version received 05.12.19; accepted 14.01.20; published 17.04.20.

Please cite as:
Stuttard L, Hewitt C, Fairhurst C, Weatherly H, Walker S, Longo F, Maddison J, Boyle P, Beresford B
Effectiveness and Cost-Effectiveness of Receiving a Hearing Dog on Mental Well-Being and Health in People With Hearing Loss: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e15452
URL: https://www.researchprotocols.org/2020/4/e15452
doi:10.2196/15452
PMID:32301737

Protocol

Delayed Auditory Feedback and Transcranial Direct Current Stimulation Treatment for the Enhancement of Speech Fluency in Adults Who Stutter: Protocol for a Randomized Controlled Trial

Narges Moein[1], MSc; Reyhane Mohamadi[1,2], PhD; Reza Rostami[3], MD; Michael Nitsche[4], MD; Reza Zomorrodi[5], PhD; Amir Ostadi[6], PhD; Abbasali Keshtkar[7], MD

[1]Department of Speech and Language Pathology, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

[2]Rehabilitation Research Center, Iran University of Medical Sciences, Tehran, Iran

[3]Faculty of Psychology, University of Tehran, Tehran, Iran

[4]Department of Psychology and Neurosciences, Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany

[5]Temerty Centre for Therapeutic Brain Intervention, Centre for Addiction and Mental Health, Department of Psychiatry, University of Toronto, Toronto, ON, Canada

[6]Faculty of Science, University of Waterloo, Waterloo, ON, Canada

[7]Department of Health Sciences Education Development, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Corresponding Author:
Reyhane Mohamadi, PhD
Department of Speech and Language Pathology
School of Rehabilitation Sciences
Iran University of Medical Sciences
Maddadkaran St, Shahid Nazari St, Mother Square, Mirdamad Blvd
Tehran, 15459-13487
Iran
Phone: 98 21 2222 2059
Email: mohamadi.r@iums.ac.ir

Abstract

Background: Stuttering is a complex speech disorder that affects speech fluency. Recently, it has been shown that noninvasive brain stimulation may be useful to enhance the results of fluency interventions in adults who stutter. Delayed auditory feedback (DAF) is a method to enhance speech fluency in individuals who stutter. Adjunctive interventions are warranted to enhance the efficacy of this intervention.

Objective: Individuals who stutter have pathological activation patterns in the primary and secondary auditory areas. Consequently, in this study, we hypothesize that stimulation of these areas might be promising as an adjunctive method to fluency training via DAF to enhance speech therapy success in individuals with a stutter. We will systematically test this hypothesis in this study.

Methods: This study is designed as a randomized, double-blind, sham-controlled clinical trial. All participants will receive DAF. The intervention group will additionally receive real transcranial direct current stimulation, while the control group will be exposed to sham stimulation. The assignment of the participants to one of these groups will be randomized. Before starting the treatment program, 2 preintervention assessments will be conducted to determine the severity of stuttering. Once these assessments are completed, each subject will participate in 6 intervention sessions. Postintervention assessments will be carried out immediately and 1 week after the last intervention session. Subsequently, to explore the long-term stability of the treatment results, the outcome parameters will be obtained in follow-up assessments 6 weeks after the treatment. The primary outcome measurement—the percentage of stuttered syllables—will be calculated in pre-, post-, and follow-up assessments; the secondary outcomes will be the scores of the following questionnaires: the Stuttering Severity Instrument–Fourth Edition and the Overall Assessment of the Speaker's Experience of Stuttering.

Results: This protocol was funded in 2019 and approved by the Research Ethics Committee of the Iran University of Medical Sciences in June 2019. Data collection started in October 2019. As of February 2020, we have enrolled 30 participants. We expect data analysis to be completed in April 2020, and results will be published in summer 2020.

Conclusions: We anticipate that this study will show an adjunctive effect of transcranial direct current stimulation, when combined with DAF, on stuttering. This should include not only a reduction in the percentage of stuttered syllables but also improved physical behavior and quality of life in adults who stutter.

Trial Registration: ClinicalTrial.gov NCT03990168; https://clinicaltrials.gov/ct2/show/NCT03990168

International Registered Report Identifier (IRRID): DERR1-10.2196/16646

(JMIR Res Protoc 2020;9(4):e16646) doi:10.2196/16646

KEYWORDS

delayed auditory feedback; stuttering; transcranial direct current stimulation; speech fluency

Introduction

Stuttering is a complex speech disorder that affects speech fluency, as defined by repetitions, prolongations, and blocks in speech sound [1]. It is prevalent in 1% of the population [2]. Despite advances in the treatment of stuttering, major limitations such as instability of treatment outcomes and lack of long-term results have yet to be addressed [3]. Recently, it has been shown that noninvasive brain stimulation may be useful to enhance the results of fluency interventions in adults who stutter [4]. One of these tools is transcranial direct current stimulation (tDCS), a noninvasive brain stimulation technique that uses a weak, constant current of 1-2 mA that passes through the scalp and results in alterations in cortical excitability and activity [5].

Structural and functional neuroimaging in individuals who stutter show bilateral underactivation of the primary and secondary auditory areas as compared to adults who do not stutter [6]. Delayed auditory feedback (DAF) is a method to enhance speech fluency in individuals who stutter. In this method, alteration of the timing feedback affects speech rate and consequently results in increased fluency [7]. A recent study has shown that in individuals who undergo DAF, enhanced speech fluency is associated with increased activity in the primary and secondary auditory areas of the superior temporal gyrus [8]. DAF is, however, time-consuming, and long-term reductions in stuttering are still restricted [9]. Adjunctive interventions are thus warranted to enhance the efficacy of this intervention. Thus, to increase speech fluency, we will combine DAF as a speech fluency intervention with tDCS applied over the superior temporal gyrus (electrode position T3 of the 10-20 international system) to enhance efficacy of the fluency intervention. Given past evidence that tDCS is capable of inducing plasticity-like changes in cortical functions that can outlast the stimulation period, we anticipate that tDCS will help stabilize intervention-related improvements in speech fluency [10].

In tDCS, electrical current flows between 2 or more electrodes—a positively charged anode and a negatively charged cathode—which are positioned at specified locations on the scalp. The current produced by tDCS results in subtle changes in the resting membrane potential of cortical neurons in the underlying brain tissue [5]. Specifically, with standard protocols, changes under the anode (referred to as anodal stimulation) result in the depolarization of critical neuronal compartments, thus increasing neuronal excitability, while changes under the cathode (cathodal stimulation) result in the hyperpolarization of respective compartments and decrease excitability at the macroscopic level [11]. Neuroplastic effects emerge after some minutes of stimulation, depending on alterations to glutamatergic and GABAergic activities [12,13]. Similar to the acute membrane polarization effects, anodal tDCS and cathodal tDCS result in excitability-enhancing and excitability-reducing plasticity, respectively [10,14,15].

In recent years, studies have shown that tDCS enhances speech fluency in adults who stutter when applied during a fluency intervention [4,16]. Since, in addition to the temporal auditory areas, the frontal cortex shows abnormalities during stuttering, tDCS was applied over the latter area in a recent study [4]. Chesters et al [4] investigated the effect of 5 sessions of anodal tDCS over the left inferior frontal gyrus during a speech fluency intervention on stuttering. Speech fluency significantly improved in the treatment group that received anodal tDCS combined with the fluency intervention in comparison with the respective sham tDCS group. It was concluded that using tDCS simultaneously with fluency training can enhance speech fluency in adults who stutter.

Due to impaired sensory-motor integration, stutterers have pathological activation patterns in the temporal lobe (ie, the primary and secondary auditory areas are underactive while speaking) [6]. Consequently, in this study, we hypothesize that stimulation of these areas might be promising as an adjunctive method to fluency training via DAF, to enhance speech therapy success in individuals with a stutter. We will systematically test this hypothesis in this study.

Methods

Overview

This project aims to investigate the effect of adjunctive noninvasive brain stimulation on speech fluency. To this aim, we will recruit 2 groups of participants—an intervention group and a control group. In the intervention group, participants will receive anodal tDCS simultaneously with DAF as a fluency intervention, whereas the participants in the control group will receive sham stimulation during DAF. The population of this study will be adults with a stutter. We hypothesize that anodal tDCS over the temporal target area will improve the efficacy of DAF to enhance speech fluency and will stabilize treatment effects.

Hypothesis

We hypothesize that the efficacy of the intervention to enhance fluency of speech in individuals with a stutter is improved when

combined with anodal tDCS over the superior temporal gyrus, as compared to a sham tDCS control condition.

Primary Objectives

The primary objective of this proposal is to compare the mean score of the percentage of stuttered syllables (SS%) [17] between the intervention and control groups at 5 time points: 1 week before treatment; immediately before treatment; and immediately, 1 week, and 6 weeks after treatment.

Secondary Objectives

The secondary objectives are as follows:

- Comparison of the mean score of stuttered syllables, duration, physical concomitant behaviors, and total score of the Stuttering Severity Instrument-Fourth Edition (SSI-4) questionnaire [18] between the intervention and control groups 1 week before treatment; immediately before treatment; and immediately, 1 week, and 6 weeks after treatment.
- Comparison of the means of the general information section, reaction to stuttering section, communication in daily situations section, quality of life section, total impact score, and impact rating of the Overall Assessment of the Speaker's Experience of Stuttering (OASES) questionnaire [19] between the intervention and control groups 1 week before treatment; immediately before treatment; and immediately, 1 week, and 6 weeks after treatment.

Efficacy

We expect adjunctive tDCS to enhance the magnitude of the effects of the intervention on our outcome parameters and to enhance the stability of improvements made due to the treatment.

Safety

There is no anticipated relevant risk associated with the stimulation parameters used in this study [20]. In a systematic review evaluating adverse effects of all published tDCS studies, including studies conducted in vulnerable populations, adverse effects of tDCS such as tingling and itching sensations under the electrodes were mild and disappeared soon after stimulation. No severe adverse effects of tDCS have been documented so far. It thus can be concluded that tDCS within the proposed stimulation parameters is safe [21].

Study Population

The study population will include adults with moderate-to-severe stuttering who have not received any treatment at least 1 month prior to the intervention. A speech-language pathologist will assess the severity of stuttering in each subject to determine whether he or she can be enrolled. The SSI-4 questionnaire will be used to assess stuttering severity. Inclusion, exclusion, and withdrawal criteria are listed in Textbox 1.

Textbox 1. Inclusion, exclusion, and withdrawal criteria.

Inclusion criteria:

- History of developmental stuttering
- Participants diagnosed with moderate-to-severe stuttering
- Right-handed
- 18 to 50 years of age (adult)
- Native speaker of Farsi
- Nonsmoker

Exclusion criteria:

- Stuttering accompanied by other speech or language disorders
- Stuttering treatment within 1 month prior to the intervention
- Hearing loss
- History of neurological or psychiatric disorders
- History of seizures
- Intake of any medication that affects brain functions (eg, antidepressants)
- Pregnancy
- Breastfeeding
- Cranial bone defects
- Cranial or brain metal implants
- Skin lesions

Withdrawal criteria:

- Request to withdraw from the study at any point during the treatment program
- Skin damage or major adverse effects of stimulation

Study Design

This study is designed as a randomized, double-blind, sham-controlled clinical trial. Participants will be randomly allocated to a control or intervention group. All participants will receive DAF during the intervention. The intervention group will additionally receive anodal tDCS, while the control group is exposed to sham tDCS. Data will be collected via questionnaires, voice recordings, and observations at 5 time points (ie, 1 week and immediately before the intervention; and immediately, 1 week, and 6 weeks post intervention). By comparing the outcomes of the respective treatments, we will investigate the efficacy of anodal tDCS combined with DAF in enhancing speech fluency.

Before starting the treatment program, 2 preintervention assessments will be conducted to determine the severity of stuttering. Once these assessments are completed, each subject will participate in 6 intervention sessions. Postintervention assessments will be carried out immediately and 1 week after the last intervention session. Afterward, to explore the long-term stability of the treatment results, the outcome parameters will be obtained in follow-up assessments 6 weeks after the treatment. This single-center trial will be conducted at the Vahdat Neurorehabilitation Clinic in Tehran, Iran.

As previously mentioned, the study will be conducted on adults with moderate-to-severe stuttering; hence, prior to recruitment, the severity of stuttering will be assessed to determine enrollment eligibility. In order to determine the severity of stuttering, an SSI-4 score for each individual will be calculated.

We will obtain two baseline measurements—one at 1 week and the other immediately before the intervention—to guarantee symptom stability. At baseline, the participants will be visited individually, and their voices will be recorded while performing 3 different tasks (ie, oral reading, monologue, and conversation). Based on these data, we will calculate the primary outcome measure (ie, SS%). The voice of the participants will be recorded by a H5 Handy Recorder ZOOM). Two experienced specialist raters will independently count the SS% of all speech samples. In order to determine interrater reliability, Cohen kappa will be calculated. In cases of agreement between the raters (intraclass correlation coefficient=0.8-1), the SS% will be reported. In cases of disagreement (intraclass correlation coefficient <0.8), a third senior rater will evaluate SS%.

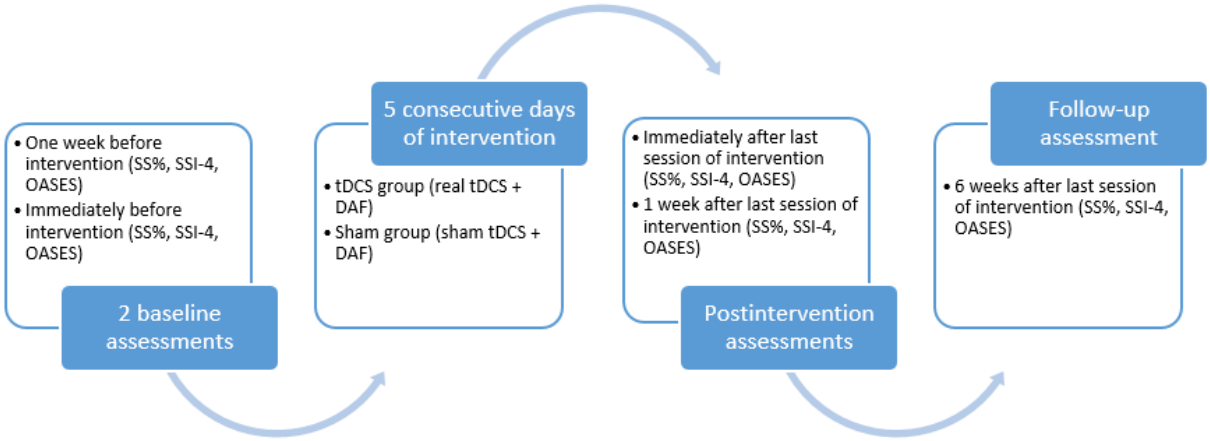
During the two baseline sessions, the principal investigator will observe physical behavior, which is required to calculate the SSI-4 score. The final part of the preintervention assessments is the completion of the OASES questionnaire by the participants. After the second baseline assessment, the intervention will start and carry on for 6 consecutive days.

At 3 time points after the last intervention session (ie, immediately, 1 week, and 6 weeks), each subject will undergo postintervention assessments, which will be identical to the preintervention assessments and obtained by the same raters. The trial design and timeline are shown in Figure 1.

After each intervention session, participants will fill out a questionnaire using a 5-point Likert scale (1=very mild, 5=very severe) to report potential side effects of tDCS, including itching, burning, tingling, headache, fatigue, sudden mood change, difficulties in concentration, changes in visual perception, unpleasant somatosensory sensations, unpleasant visual sensations, nausea, drowsiness, persisting feelings of stimulation, and 1 open question for any other adverse effects.

In addition, to ensure successful blinding, participants will be asked to guess the type of treatment they will or had received (ie, anodal tDCS or sham) before and after the intervention.

Figure 1. Trial design and timeline. DAF: Delayed Auditory Feedback; OASES: Overall Assessment of the Speaker's Experience of Stuttering; SS%: percentage of stuttered syllables; SSI-4: Stuttering Severity Instrument-Fourth Edition; tDCS: transcranial direct current stimulation.



Randomization and Blinding

This study is a randomized, double-blind, sham-controlled trial. Participants will be randomly allocated to the intervention and control groups. The random assignment of participants ensures the prevention of a possible selection bias and a disbalance of confounding factors between the study arms. Randomization will be performed via a web-based randomization tool [22]. In accordance with the requirements of this tool, we will include 50 subjects and 4 blocks of equal size, for which randomization will be performed independently. Each subject will be given an ID number and will be assigned to one of the treatment groups. The individual responsible for generating the random list will not be involved in any other part of the trial.

As the study is double-blinded, neither the participants nor the investigators will know which group each participant will be assigned to. Sealed opaque envelopes will be used for concealment. Accordingly, an envelope will be produced for each subject. These envelopes will be marked with an ID number, and the treatment group associated with each ID will be placed in its respective envelope. In order to guarantee double blinding, specific letters, unknown to the principal investigator and the subjects, will be used instead of the actual names of the treatment groups (ie, intervention and control) inside the envelopes. Assignment of these letters to the respective intervention group will be done by a clinician (not the principal investigator). Before the principal investigator starts the treatment of the respective participant, the subject will be given one of the sealed envelopes by the clinician. Once the envelope is opened by the corresponding participant, the clinician will set up the mode—anodal or sham—based on the content of the subject's envelope and attach the electrodes to the participant´s head as illustrated in the sections below. Once completed, the principal investigator will commence the treatment. The researchers who are responsible for outcome assessment and data analysis will also be blinded to the intervention groups.

Intervention Program

Performance of Transcranial Current Direct Stimulation

The stimulation will be done by passing a current of 1 mA between two 5 cm × 7 cm electrodes for a duration of 20 minutes with ramp-up and ramp-down intervals of 15 seconds. This common setup has been shown to be efficient for enhancing speech fluency in previous tDCS studies [3,4]. A neuroConn DC-STIMULATOR will be used to deliver tDCS. The electrodes used for tDCS are conductive rubber electrodes encased in a sponge pocket, and saline solution will be used as an electrolyte-based contact medium. The electrode sponges must be saturated with saline. It is crucial that each side of the sponges is sufficiently moisturized, but not overly saturated, to avoid dripping. Before placing electrodes on the scalp, the clinician will inspect the skin for any skin damage or lesion. For both anodal and sham modes, the same stimulation intensity parameters will be used. However, for the sham stimulation, the device will ramp down automatically after 30 seconds. The intervention will be conducted for 6 consecutive days.

An electroencephalogram cap will be used to identify the site of stimulation. The anode electrode will be placed over the left superior temporal gyrus (T3 according to the 10-20 international system [23]), and the cathode electrode will be positioned over the right frontopolar region (Fp2 according to the 10-20 system [23]). To prevent movement of the electrodes, their positions will be fixed by elastic rubber straps. Electrode placement is shown in Figure 2.

Figure 2. Electrode placement.

Performance of Delayed Auditory Feedback

During stimulation, participants of both groups will receive DAF as a fluency treatment. In order to deliver DAF, Audapter, which is a software package for manipulating the acoustic parameters of speech in real time, will be used [24]. This package consists of the core algorithm for real-time manipulation and a MATLAB wrap-around. The real-time signal processing algorithms are coded in C++. The subjects will perform 3 tasks (ie, oral reading, monologue, and conversation) with DAF <60 ms. This 60-ms delay has been shown to be efficient to enhance speech fluency in individuals with a stutter [25].

Although, as outlined above, we do not anticipate major adverse events, the occurrence of skin irritations, itching, tingling, burning, or pain (including headache) will be monitored. After each intervention session, a side effects questionnaire will be filled out by participants. Furthermore, participants will guess if they received anodal or sham stimulation before the intervention, after the first session, and after the intervention period to ensure successful blinding.

Data from the respective questionnaires will be recorded on paper, and recordings of subjects' voices will be stored electronically. Access to data resources will be restricted to the investigators. Finally, all study documents will be securely maintained for 2 years.

Statistical and Analytical Analyses

In this proposal, the null hypothesis (H_0) is that the combined treatment method using tDCS stimulation and DAF, as compared to DAF combined with sham stimulation, has no additional effect on the enhancement of speech fluency in adults who stutter.

Intended Sample Size

Based on the previous study by Chester et al [4], and using G*Power software [26], 25 subjects per group (50 participants in total) are required to detect a significant difference for a time × group interaction using a two-way mixed model analysis of variance (ANOVA). Time is the within-subject factor (ie, before the intervention and immediately, 1 week, and 6 weeks after the intervention), and group is the between-subject factor (ie, anodal or sham tDCS). The dependent variable is stuttered syllables (SS%). Type I (α) and type II (β) errors are set at .05 and .20, respectively, and the effect size is 0.17.

Other Statistical Analyses

Descriptive statistics will be calculated for demographic and baseline characteristics as well as primary and secondary outcomes for all participants. For every subject, summary statistics including mean, SD, median, minimum, and maximum will be provided for quantitative variables. For qualitative variables, frequency tables will be presented. Group differences of demographic variables will be assessed by Student *t* tests for quantitative variables and chi-square tests for qualitative variables.

In order to evaluate the effect of the treatment, the early and late outcomes (ie, measurements taken immediately, 1 week after treatment, and 6 weeks after treatment) will be examined. Depending on the significance of the ANOVA results, exploratory post hoc Student *t* tests will be conducted to compare conditions. The same procedures will be conducted for the secondary outcome parameters.

Data on tolerability and safety will be analyzed using an independent samples *t* test for common adverse events. Reports of rare side effects will be documented. To confirm successful blinding, a chi-square test will be applied.

Participants may be lost to follow-up for two reasons: (1) they have not completed at least 80% of the treatment (5 out of 6 sessions) and (2) they did not show up for the outcome measurement sessions 1 week or 6 weeks after the intervention. In these cases, intention-to-treat and modified intention-to-treat approaches will be employed. For cases lost to follow-up, efforts will be made to identify the reason, for example, by phone.

Ethics Approval and Consent to Participate

This study will be carried out in accordance to the ethical principles and national norms and standards for conducting medical research in Iran (approval ID: IR.IUMS.REC.1398.352, Iran National Committee for Ethics in Biomedical Research). In addition, our informed consent form will be reviewed and approved by the ethics committee. Prior to the commencement of any part of the study, the principal investigator will obtain written informed consent from all participants.

Consent for Publication

All study findings and information will be posted on ClinicalTrials.gov (identifier NCT03990168). Although the investigator will be free to use the study findings for educational and scientific purposes, written consent will be obtained from the study sponsor prior to submission of any manuscript for publication.

Availability of Data and Materials

Information about subjects will be kept confidential and will not be available for public access. In addition, any document, data, voice recording, and other records will be identified by a participant ID number, and the name of each subject will be kept confidential. The data sets used and analyzed during the study are available from the corresponding author upon reasonable request.

Results

This protocol was funded in 2019 and approved by the Research Ethics Committee of the Iran University of Medical Sciences in June 2019. Data collection started in October 2019. As of February 2020, we have enrolled 30 participants. We expect that data analysis will be completed in April 2020, and the results will be published in summer 2020.

Discussion

This is the first randomized controlled trial with follow-up measures to explore the efficacy of tDCS combined with DAF to improve speech fluency in adults who stutter. In this study, the impact of the combined treatment will be evaluated by SS% as the primary outcome measure. Treatment impact will be examined for up to 6 weeks to exclude any temporary effects with limited clinical value [1]. The findings of this study will thus supply information about the efficacy of tDCS as an adjunctive therapy for reducing SS% and its stability over a prolonged time period.

In addition, we will obtain the SSI-4 score before and after treatment. This score includes additional information such as changes in the duration of stutter moments and physical concomitants of stuttering. These parameters disrupt speech fluency and have negative effects on communication. Reduction in these behaviors, as expected, are relevant for the enhancement of communication efficacy in adults who stutter [18].

Finally, we will survey the impact of stuttering on a person's quality of life by using the OASES questionnaire. As a comprehensive assessment of stuttering, this questionnaire measures the effect of stuttering on multiple life situations [19].

In summary, we anticipate that this study will show an adjunctive effect of tDCS, when combined with DAF, on stuttering. This should include not only a reduction in SS% but also improved physical behavior and quality of life in adults who stutter.

Acknowledgments

We would like to thank the Iran National Science Foundation for supporting this trial. The Iran University of Medical Sciences will fund this work. The funding body has a role in the design of the study; collection, analysis, and interpretation of data; and writing of the manuscript.

Authors' Contributions

RM, RR, and MN have contributed to the conception and design of the study. AK and RZ will contribute to data analysis and interpretation. AO has contributed to writing the MATLAB code. NM was a major contributor in writing the manuscript and will collect the data. MN substantively revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review comments from the Iran University of Medical Sciences.
[PDF File (Adobe PDF File), 12726 KB - resprot_v9i4e16646_app1.pdf]

References

1. Guitar B. Stuttering: An integrated approach to its nature and treatment. Baltimore, MD: Lippincott Williams & Wilkins; 2013.
2. Bloodstein O. A handbook on stuttering. Chicago, IL: Singular Publishing Group; 1995.
3. Chesters J, Watkins KE, Möttönen R. Investigating the feasibility of using transcranial direct current stimulation to enhance fluency in people who stutter. Brain and Language 2017 Jan;164:68-76. [doi: 10.1016/j.bandl.2016.10.003]
4. Chesters J, Möttönen R, Watkins KE. Transcranial direct current stimulation over left inferior frontal cortex improves speech fluency in adults who stutter. Brain 2018 Apr 01;141(4):1161-1171 [FREE Full text] [doi: 10.1093/brain/awy011] [Medline: 29394325]
5. Thair H, Holloway AL, Newport R, Smith AD. Transcranial Direct Current Stimulation (tDCS): A Beginner's Guide for Design and Implementation. Front Neurosci 2017 Nov 22;11:641 [FREE Full text] [doi: 10.3389/fnins.2017.00641] [Medline: 29213226]
6. Brown S, Ingham RJ, Ingham JC, Laird AR, Fox PT. Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. Hum Brain Mapp 2005 May;25(1):105-117. [doi: 10.1002/hbm.20140] [Medline: 15846815]
7. Howell P. Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering. Stammering Res 2004 Apr 01;1(1):31-46 [FREE Full text] [Medline: 18259594]
8. Takaso H, Eisner F, Wise RJS, Scott SK. The Effect of Delayed Auditory Feedback on Activity in the Temporal Lobe While Speaking: A Positron Emission Tomography Study. J Speech Lang Hear Res 2010 Apr;53(2):226-236. [doi: 10.1044/1092-4388(2009/09-0009)]
9. Van Borsel J, Reunes G, Van den Bergh N. Delayed auditory feedback in the treatment of stuttering: clients as consumers. Int J Lang Commun Disord 2003 Jan;38(2):119-129. [doi: 10.1080/1368282021000042902]
10. Nitsche MA, Paulus W. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. Neurology 2001 Nov 27;57(10):1899-1901. [doi: 10.1212/wnl.57.10.1899]
11. Nitsche MA, Schauenburg A, Lang N, Liebetanz D, Exner C, Paulus W, et al. Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human. J Cogn Neurosci 2003 May 15;15(4):619-626. [doi: 10.1162/089892903321662994] [Medline: 12803972]
12. Nitsche M, Fricke K, Henschke U, Schlitterlau A, Liebetanz D, Lang N, et al. Pharmacological modulation of cortical excitability shifts induced by transcranial direct current stimulation in humans. J Physiol 2003 Nov 15;553(Pt 1):293-301 [FREE Full text] [doi: 10.1113/jphysiol.2003.049916] [Medline: 12949224]
13. Stagg CJ, Best JG, Stephenson MC, O'Shea J, Wylezinska M, Kincses ZT, et al. Polarity-Sensitive Modulation of Cortical Neurotransmitters by Transcranial Stimulation. Journal of Neuroscience 2009 Apr 22;29(16):5202-5206. [doi: 10.1523/jneurosci.4432-08.2009]
14. Nitsche MA, Nitsche MS, Klein CC, Tergau F, Rothwell JC, Paulus W. Level of action of cathodal DC polarisation induced inhibition of the human motor cortex. Clinical Neurophysiology 2003 Apr;114(4):600-604. [doi: 10.1016/s1388-2457(02)00412-1]
15. Nitsche M, Paulus W. Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. J Physiol 2000 Sep 15;527 Pt 3(3):633-639 [FREE Full text] [doi: 10.1111/j.1469-7793.2000.t01-1-00633.x] [Medline: 10990547]
16. Chesters J. Enhancing speech fluency using transcranial direct current stimulation. Oxford, UK: University of Oxford; 2016.
17. Cordes AK, Ingham RJ. The Reliability of Observational Data. J Speech Lang Hear Res 1994 Apr;37(2):279-294. [doi: 10.1044/jshr.3702.279]
18. Riley G. SSI-4 stuttering severity instrument fourth edition. Austin, TX: Pro-Ed; 2009.
19. Yaruss JS, Quesal RW. Overall Assessment of the Speaker's Experience of Stuttering (OASES): Documenting multiple outcomes in stuttering treatment. Journal of Fluency Disorders 2006 Jan;31(2):90-115. [doi: 10.1016/j.jfludis.2006.02.002]
20. Nitsche MA, Liebetanz D, Lang N, Antal A, Tergau F, Paulus W. Safety criteria for transcranial direct current stimulation (tDCS) in humans. Clinical Neurophysiology 2003 Nov;114(11):2220-2222. [doi: 10.1016/s1388-2457(03)00235-9]
21. Matsumoto H, Ugawa Y. Adverse events of tDCS and tACS: A review. Clinical Neurophysiology Practice 2017;2:19-25. [doi: 10.1016/j.cnp.2016.12.003]
22. Randomization.com. 2017. URL: http://www.randomization.com/ [accessed 2020-03-18]
23. Jasper, H. Report of the committee on methods of clinical examination in electroencephalography. Electroencephalography and Clinical Neurophysiology 1958 May;10(2):370-375. [doi: 10.1016/0013-4694(58)90053-1]
24. Cai S. Tensorflow debugger: Debugging dataflow graphs for machine learning. Proceedings of the Reliable Machine Learning in the Wild - NIPS 2016 Workshop 2016:2016.
25. Stuart A, Kalinowski J, Rastatter MP, Saltuklaroglu T, Dayalu V. Investigations of the impact of altered auditory feedback in ‐ the ‐ ear devices on the speech of people who stutter: initial fitting and 4 ‐ month follow ‐ up. Int J Lang Commun Disord 2004 Jan;39(1):93-113. [doi: 10.1080/13682820310001616976]

26. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/brm.41.4.1149]

Abbreviations

ANOVA: analysis of variance
DAF: delayed auditory feedback
OASES: Overall Assessment of the Speaker's Experience of Stuttering
SS%: percentage of stuttered syllables
SSI-4: Stuttering Severity Instrument-Fourth Edition
tDCS: transcranial direct current stimulation

Edited by G Eysenbach; submitted 10.10.19; peer-reviewed by S D'Arcy, E Upton; comments to author 18.02.20; revised version received 25.02.20; accepted 27.02.20; published 21.04.20.

Please cite as:
Moein N, Mohamadi R, Rostami R, Nitsche M, Zomorrodi R, Ostadi A, Keshtkar A
Delayed Auditory Feedback and Transcranial Direct Current Stimulation Treatment for the Enhancement of Speech Fluency in Adults Who Stutter: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e16646
URL: http://www.researchprotocols.org/2020/4/e16646/
doi:10.2196/16646
PMID:32314973

Protocol

Effects of Sharing Old Pictures With Grandchildren on Intergenerational Relationships: Protocol for a Randomized Controlled Trial

Zoljargalan Gantumur[1], MSW; Marcos Baez[2,3], PhD; Nomin-Erdene Ulamnemekh[1], BSc; Francisco Ibarra[2], PhD; Sugarmaa Myagmarjav[1], PhD; Fabio Casati[2,4], PhD

[1]School of Public Health, Mongolian National University of Medical Sciences, Ulaanbaatar, Mongolia

[2]University of Trento, Povo (TN), Italy

[3]Université Claude Bernard Lyon 1, Lyon, France

[4]School of Engineering Entrepreneurship, Tomsk Polytechnic University, Tomsk, Russian Federation

Corresponding Author:
Marcos Baez, PhD
University of Trento
Via Sommarive, 9
Povo (TN)
Italy
Phone: 39 0461 28 3966
Email: baez@disi.unitn.it

Abstract

Background: Intergenerational relationships are beneficial for both grandparents and grandchildren. A positive grandparent-grandchild relationship can improve the psychological well-being of older adults and be a source of social support, family history, and identity development. Maintaining meaningful interactions can be, however, a challenging endeavor, especially as life events lead to relocating geographically. Grandparents and grandchildren can have different preferences in terms of communication mediums and different assumptions about the real conversational needs of the other.

Objective: In this study, we will investigate the feasibility and effect of sharing memories of older adults with their grandchildren in social media. This intervention focuses on bringing snippets of the lives of the grandparents into the grandchildren's social media feed and analyzing the potential effect on relational quality, relational investment, and conversational resources from the perspective of the grandchildren.

Methods: A randomized controlled trial will be used to measure the effectiveness of sharing family memories through social media on intergenerational relationships from the perspective of the grandchildren. The study will be implemented in Mongolia among 60 grandparent-grandchild pairs who will be assigned to either a control or intervention group. Pictures and stories will be collected during reminiscence sessions between the researchers and the grandparents before the intervention. During an intervention period of 2 months, grandchildren in the intervention group will receive pictures and stories of their grandparents on their social media account. Pre- and postintervention questionnaires will measure relationship quality, relationship investment, and conversational resources and will be used to assess the effectiveness of the intervention.

Results: We conducted a pretest pilot from January to April 2018 among 6 pairs of participants (6 grandparents and 6 grandchildren). The validation of the protocol was focused on the process, instruments, and technological setup. We continued the study after the validation, and 59 pairs of participants (59 grandparents and 59 grandchildren) have been recruited. The data collection was completed in November 2019.

Conclusions: The results of this study will contribute to strategies to stimulate social interactions in intergenerational pairs. A validation of the study process is also presented to provide further operational recommendations. The lessons learned during the validation of the protocol are discussed with recommendations and implications for the recruitment, reminiscence sessions, technological setup, and administration of instruments.

International Registered Report Identifier (IRRID): DERR1-10.2196/16315

(JMIR Res Protoc 2020;9(4):e16315) doi:10.2196/16315

KEYWORDS

intergenerational relationship; social media; Facebook intervention; closeness

Introduction

Importance of Intergenerational Interactions

Vast literature in familial intergenerational interactions is devoted to the study of the grandparent-grandchild relationship. Older adults not only benefit with joy, love, closeness, and company from grandchildren [1], but the expression of affection toward grandchildren is positively associated with psychological well-being (eg, reduction in loneliness and stress and increased general mental well-being) [2]. For younger adults, interactions with grandparents account for the majority of intergenerational interactions [3]. Grandchildren benefit from these relationships affectively, cognitively, and materially [1]. Grandparents are a source of social support, family history and identity, and development [4,5]. However, maintaining and developing intergenerational interactions is a challenging endeavor and there are barriers to interactions.

Barriers

Geographic Separation

Geographic distance is one of the turning points in intergenerational interactions and particularly in the grandparent-grandchild relationship [6], as it limits opportunities for face-to-face interactions. However, studies have shown that grandparents can still experience high levels of relational quality using other means of communication (eg, phone or email) if communication is frequent and especially if initiated not only by the grandparents [7,8].

Scant Use of Conversational Resources and Skills

Intergenerational conversations can be challenging, as interlocutors have lived through different life periods and might have different assumptions, conversational skills, and needs [9], which may affect the ability to find and engage on appropriate conversation topics for both. On the ability to engage, the perception of properly accommodating the communication to the other person in the conversation is a predictor of communication satisfaction, liking, and emotional closeness [10]. Failing to accommodate and attune the communication to the interlocutor, however, can lead to patronizing speech, painful disclosures, or underutilization of topical resources [9] that affect the satisfaction and willingness to participate in such interactions [11]. As for finding appropriate topics, the lack of conversation topics is reported as a source of anxiety in intergenerational interactions [12].

Asymmetry in Relational Investment

Lack of relational investment is yet another challenge in an intergenerational relationship [6]. It particularly affects young adults transitioning to adulthood, a period in which the responsibilities of growing older are associated with changes in levels of interactions and relational closeness [13]. Grandparents, on the other hand, do keep an interest in their grandchildren's life, but often refrain from contact, asking questions, or addressing some topics, not wanting to be an annoyance and thus relying on the social skills of the young [14]. Social media have opened an opportunity for older adults to learn more about their grandchildren. However, interactions in this medium are most often not reciprocated by older adults [15].

Objectives

Interventions to promote social interactions have the potential to address challenges. Systematic reviews summarizing years of research show that computer and internet interventions can be effective in reducing loneliness among older adults [16-18]. However, very few interventions have focused on the specific challenges of intergenerational interactions, especially from the side of the young relatives as they navigate periods of complex social changes. Previous studies on technology-mediated intergenerational communication made different assumptions regarding the younger generation's relationship with their grandparent and their relational needs: (1) some studies assume that the younger relative is already invested in the relationship and actively seeks contact [19], (2) other studies focus on the opposite problem, that of making the older family members aware of the activities of their younger family members [20,21], and (3) a third body of literature aims to bring older adults online and make online interactions with younger adults more reciprocal [22]. This calls for more research into intervention strategies to improve the relational quality in the grandchild-grandparent relationship by focusing on the challenges faced by the younger population.

In this paper, we present a protocol for a study that explores the effect on relational quality of selectively sharing grandparents' life stories with grandchildren on social media. We focus on the effect that such intervention can have on grandchildren, and particularly young adults, as changing their communication behavior and investment can have great effects on the relationship. We know from previous work that relationships where grandchildren and grandparents are equally likely to initiate conversations are perceived as more satisfying in terms of frequency and quality by grandparents [7]. A previous study also reported that grandparents tend to rate the relationship as very close while their grandchildren as not having a close connection, suggesting different generational perspectives [23]. Thus, in a period of their lives marked by many social changes, approaches to maintaining or increasing frequency and quality social contact with grandparents and the sense of closeness is of paramount importance.

The sharing of family memories in this context is unidirectional: family memories are used as triggers to give grandchildren topics of conversations, stimulate social interactions, and eventually increase their sense of connectedness. All social interactions with grandparents happen outside social media, either by phone or face to face, and, in principle, without the grandparents having to use technology or have a social media presence. This approach is well suited to scenarios where the practice of reminiscence is widely adopted, such as in nursing homes, where family memories can be used as a bridge to

connect grandchildren with their (often frail) grandparents [24,25].

The study is inspired by the practice of reminiscence, the process of recollecting past memories—a practice that is common at all ages [26] and often conducted with older adults due to its many functions and benefits. Webster [27] identifies 8 particular functions: death preparation, identity, problem solving, teach and inform, conversation, boredom reduction, bitterness revival, and intimacy maintenance. Thus, reminiscence serves an important social function in facilitating the sharing of personal memories with others, helping to create bonds between people [28], offering an interesting platform for grandchildren to learn about their grandparents and build family history [5], potentially reconnecting with their grandparents.

In this protocol, we focus on studying the effects of sharing grandparents' pictures and related stories with grandchildren on social media. Thus, the sharing is mediated by technology. In doing so, we address the following main research question:

- RQ1: Does sharing old pictures and related short stories (from grandparents) with young relatives increase their relational quality?

With this research question, we study the impact of the intervention on the intergenerational relational quality from the perspective of the grandchildren, measured as the feeling of closeness with their grandparents. In addition, we look at two main factors contributing to relational quality and observe if these factors increase as a result of the intervention. More precisely, we address the following secondary research question:

- RQ2: Does sharing old pictures (from grandparents) with the young relatives increase their relational investment?

We particularly focus on relational investment as captured by the impact on the grandchildren's commitment in their relationship with their grandparents, as well as on the frequency of communication.

- RQ3: Does sharing old pictures (from grandparents) with the young relatives provide them with more topics of conversations and relationship chemistry?

We focus on conversational resources as captured by the impact on the openness and predisposition toward each other (relationship chemistry), as well as on the number and diversity of topics of conversation.

The potential impact of the study is to validate an approach to improving relationship quality between grandparents and grandchildren by empowering grandchildren through the sharing of family stories. This approach can complement existing online and colocated reminiscence technology that generally assume active collaboration of family members to keep the older relatives socially active.

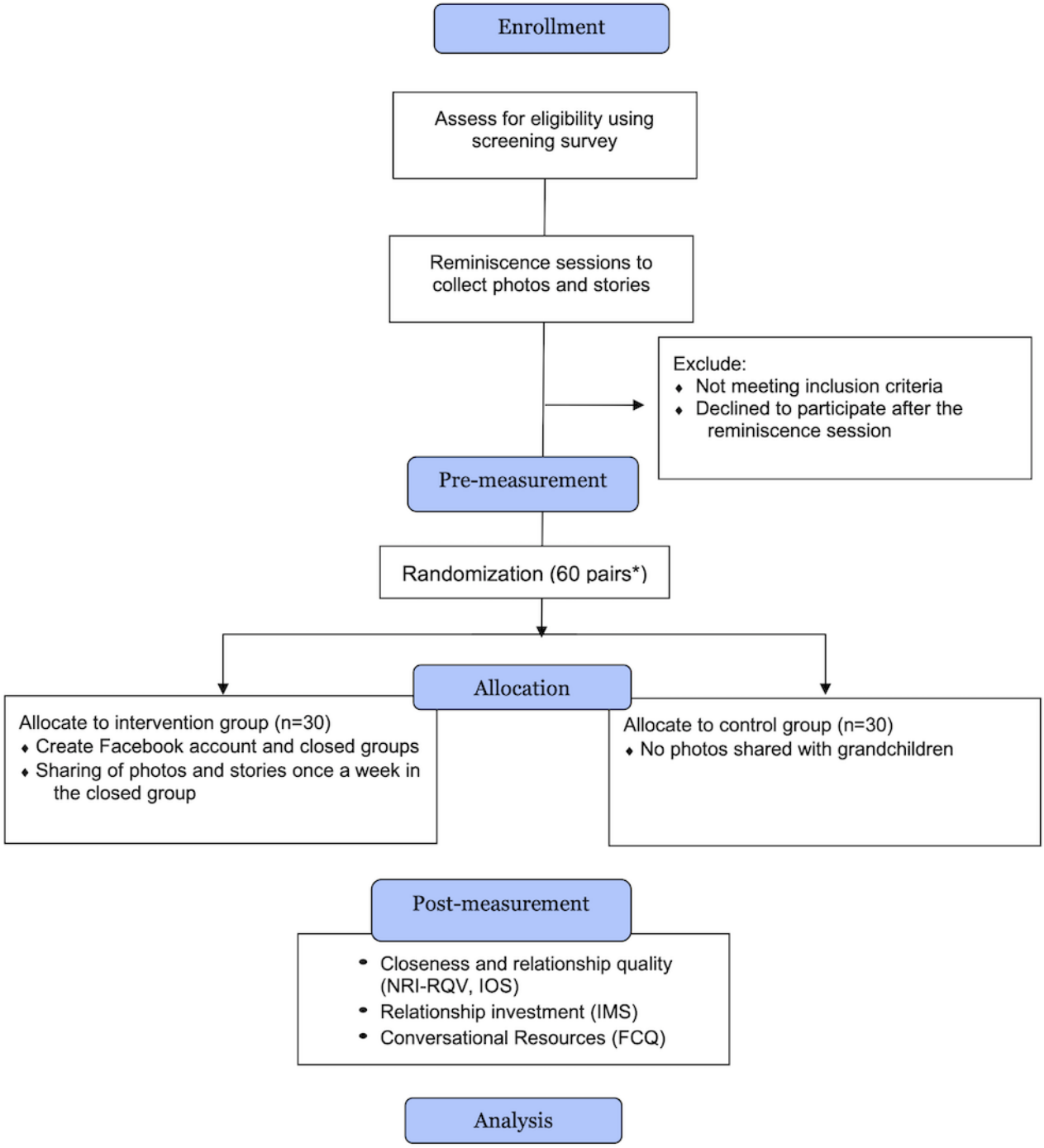
Methods

Study Design

The study is designed as a randomized controlled trial (RCT) where both the intervention and control group consist of grandparent-grandchildren pairs. In the intervention group, we collect a set of pictures and short stories about the grandparents and share them with the grandchildren's social network account, one at a time, at regular intervals across a period of 2 months. In the control group, there is no sharing or collecting pictures as the goal is to account for alternative effects.

Figure 1 shows the diagram of the study flow. Restricted randomization will be performed by stratifying gender, initial closeness level of the grandchildren, and age group of the grandparents. This paper reports on a version of the protocol accepted by the Research Ethics Committee at Mongolian National University of Medical Sciences in December 2017 (No. 2017/3-05).

Figure 1. Study Flow. *: 60 grandchildren will be paired with their grandpaerents; NRI-RQV: Network of Relationships Inventory–Relationship Quality Version; IOS: Inclusion of Other in the Self scale; IMS: Investment Model Scale; FCQ: Friendship Chemistry Questionnaire.



Participants

The aim is to recruit 60 grandparent-grandchild pairs, which will be split evenly into intervention and control conditions. The sample size is limited by the specific focus of the intervention and the availability of participants complying with the criteria, but it represents a valid sample for the statistical analyses. A convenience sample will be used in this study to recruit participants. We consider eligible pairs of grandchildren and grandparents meeting the following inclusion criteria.

Grandchildren should be willing to participate in the study and meet the following requirements:

- Aged 18 to 30 years: we recruit young adults because this is a transition period to adulthood associated with changes in the level of interactions and relational closeness [13]

- Be regular users of Facebook: we use Facebook because this is the most common social media among Mongolian young adults [29], and we define regular users as those who check their accounts every 2 to 3 days or more often
- Have infrequent meetings with their grandparent: the proposed intervention is aimed at improving relationships between grandparents and grandchildren who do not have much opportunity to meet and talk. Therefore, we recruit grandchildren who live separately from their grandparents (not in the same apartment or building) and do not meet with their grandparents more than once a week on average

In addition, the grandchildren should have one grandparent willing to participate in the study who meets the following requirements:

- Retired and aged over 60 years: according to Mongolian law, the retirement age is 55 years for women and 60 years for men
- Be able to independently communicate with a researcher: the selected grandparent will be requested to share their life story and give permission to researchers to post photos on Facebook. Therefore, we recruit grandparents who are competent to participate in the study
- Not have any account on social media: the proposed intervention aims at sharing pictures and related stories with the grandchildren in a controlled environment, and thus assumes there is no other form of sharing in the grandparent-grandchild pairs occurring on social networks. Therefore, including grandparents who use social media will not match the study objective
- Have at least 15 available pictures about their life: during the intervention period, we will share Facebook posts 8 times (once a week). Each Facebook post will consist of at least 1 or 2 photos that illustrate a story. In addition, the stories will be about the different life span of the grandparent; therefore, we need 15 photos to illustrate different life stories of the grandparent

Recruitment

Grandchildren will be recruited from the public universities in Ulaanbaatar, Mongolia, using online as well as paper-based recruitment surveys. As a first step, we will reach students by introducing the study and its goals at the end of their lectures. Those interested in participating in the study will then stay in the classroom to complete the recruitment survey and sign the consent form. The survey will include filtering questions, reflecting the inclusion criteria for grandchildren, and will ask participants to provide an email or phone number for a follow-up. Students who met our inclusion criteria and who have stated their willingness to invite their grandparents to join in our study will be contacted by the researchers to confirm the interest of their grandparents and organize a phone interview. The list of potential candidates will be managed electronically, with access provided only to the researchers in charge of the recruitment.

As a second step, the researchers will contact the grandparents for a phone interview to confirm their eligibility. At the end of the phone interview, a follow-up visit will be scheduled with the grandparents who met the inclusion criteria. We will recruit study participants and make visits until we have 60 pairs that meet the inclusion criteria and agree to participate.

Processes and Intervention

Home Visit to Collect Photos and Stories

Researchers will make a home visit according to the scheduled appointments with the selected grandparents. At the beginning of the meeting, researchers will introduce themselves, explain the purpose of the study, and request that the grandparent read and sign the consent form. Researchers will then ask grandparents to bring pictures that illustrate their life history that they would be willing to share with their grandchildren. In order to ensure diversity in the pool of pictures to share, researchers will propose selecting pictures with the following characteristics: (1) relevant life events such as birth, marriages, work, etc; (2) pictures the grandparents like, find funny or interesting, or that they feel proud of; (3) pictures that include children as well as grandchildren; and (4) pictures that are very old and some more recent. During the visit, researchers will start a conversation regarding the photos by asking the story behind each picture. The researchers will digitize the picture (eg, taking a picture with a mobile phone) and take note of the stories told during the reminiscence session.

Premeasurement

Researchers will schedule a meeting with the participating grandchildren to complete the premeasurement questionnaire, which includes all instruments described in the Outcome Measures section.

Intervention

The intervention will be implemented over a period of 2 months among the participants of the intervention group. The sharing of pictures and stories will be done using Facebook closed groups. A closed group will be created for each grandchild with whom posts will be shared privately by one of the researchers using an account created for the sole purpose of the study. Grandchildren will be invited to their specific closed group before the start of the intervention with a welcome post explaining in detail the procedure of the picture sharing. A privacy feature of this setup is that sharing outside closed groups is not enabled, which means that grandchildren will not be able to share the original posts outside the closed group.

During the intervention, grandchildren will receive one post per week. The post will be composed by the research team using the pictures and stories collected during the reminiscence session. Each post will include a picture and the related story written as told by the grandparent to the researcher. It is important to note that the older adults will not be directly involved in the intervention process, as this intervention does not enable direct communication between grandparents and grandchildren (ie, grandchildren will not be able to interact with their grandparents through Facebook). Any potential interaction in the grandparent-grandchild pairs will occur through other mediums. Participants in the control condition will only participate in the pre- and postmeasurement sessions. All participants are informed verbally and in the informed consent of their rights, which includes leaving the study at any moment.

Postmeasurement

After 2 months of intervention, the grandchildren in both intervention and control groups will be invited for an individual meeting to perform postmeasurement. The same questionnaires provided in the premeasurement session will be used. In addition, the participating grandchildren will be debriefed on their experience during the intervention, and on if and how it affected their relationship with their grandparent.

Outcome Measures

Primary Outcomes

The Inclusion of Other in the Self (IOS) scale and the Network of Relationships Inventory–Relationship Quality Version (NRI-RQV) will be used to measure the relationship of grandparent and grandchild.

IOS is a pictorial tool developed to measure the degree of interpersonal connectedness. In this study, a modified version of the IOS by Gächter et al [30] will be used. This instrument was chosen because it is validated and more comprehensible than the original measure. The NRI-RQV is a combination of the Network of Relationships Inventory and a family relationship measure developed by Buhrmester and Furman [31]. This 30-item survey has 10 scales with 3 items per scale. In this study, the 5 positive features—companionship, intimate disclosure, satisfaction, emotional support, and approval—of the NRI-RQV will be used as a measure of closeness between grandchildren and grandparents. The original questions were adapted by replacing the pronouns with grandparent to fit the grandparent-grandchild relationship.

Secondary Outcomes

We will rely on the Investment Model Scale (IMS) [32] and self-reported communication frequency to measure relationship investment. The IMS is an instrument developed to measure commitment level and 3 bases of dependence, which are satisfaction level, quality of alternatives, and investment size. The IMS includes a specific subscale on relationship investment that will be used to measure the grandchild's contribution in this relationship. The original questions were adapted to the grandparent-grandchild relationship by replacing the word partner with grandparent. Relationship investment scales are scored by computing the scores with higher scores representing higher commitment. Communication frequency between the grandparent and grandchildren will be measured with a custom question with response values going from a few times a week to never.

We rely on the Friendship Chemistry Questionnaire (FCQ) [33] and a custom-made set of questions to measure conversational resources. The FCQ was developed to explore friendship formation factors, comprising 5 subscales: reciprocal candor, mutual interest, personableness, similarity, and physical attraction [33]. In this study, reciprocal candor and similarity scales of the FCQ will be used to estimate available conversational resources, as they reflect on the openness and predisposition toward each other. The original questions were adapted from friendship to the grandparent-grandchild relationship by replacing the pronoun. The FCQ is scored by averaging the items with high scores representing the higher levels of each scale. Two separate questions were redacted and added by researchers to assess general conversational resources and identify common conversational topics with grandparents.

As part of the debriefing interview with participants, researchers will inquire about how the intervention affected their relationship with their grandparent, including recent positive and negative episodes of interaction with grandparents and what pictures and stories captured their attention and led to more interactions. These insights will be contrasted with the online reactions and comments on the posts.

Statistical Analysis

We will analyze relational quality and closeness (RQ1) with analysis of variance (ANOVA), with group (control, intervention) as between-subject factor and time (pre- and postmeasurement) as within-subject factor. We will compute the main effect for time and the interaction between time and group. In addition, we will analyze the correlation between connectedness and closeness and the factors contributing to higher increases in relational quality, if any.

To determine a statistically significant difference in the relational investment score (RQ2) between control and intervention groups as a result of the intervention, we will perform an ANOVA with group (control, intervention) as between-subject factor and time (pre- and postmeasurement) as within-subject factor. We will compute the main effect for time and the interaction between time and group. In addition, we will explore the factors contributing to higher increases in relational investment, if any.

To determine a statistically significant difference in the conversational resources score (RQ3) between control and intervention groups as a result of the intervention, we will perform an ANOVA with group (control, intervention) as between-subject factor and time (pre- and postmeasurement) as within-subject factor. We will compute the main effect for time and the interaction between time and group. In addition, we will explore the factors contributing to higher increases in conversational resources, if any. The resulting analysis will include an analysis of any unexpected or adverse effects of the intervention.

Results

Pilot

In order to validate the protocol and refine the entire intervention process by collecting actionable recommendations for how to execute the various phases, we conducted a pretest pilot from January to April 2018 among 6 pairs of participants (6 grandparents and 6 grandchildren). The validation of the protocol was focused on the process, instruments, and technological setup. We continued the study after the validation, and 59 pairs of participants (59 grandparents and 59 grandchildren) were recruited. Data collection was completed in November 2019.

Challenges and Lessons Learned

Low Recruitment Rate of Study Participants

We recruited study participants at the Mongolian National University of Medical Sciences. We screened approximately 260 students to recruit participants who met the inclusion criteria. Students were excluded because the inclusion criteria "living separately from their grandparents" and "grandparents living in Ulaanbaatar city" were not met. The low rate of recruitment was related to the fact that the number of students who came from the countryside is much higher than that of students who reside in the city. Based on this experience, we recommend extending the recruitment pool to reach young participants who meet the age criteria but are not necessarily affiliated with any university. In addition, disseminating the online recruitment survey is another possible solution to increase recruitment rate.

Scheduling Issues in Home Visits

Although older adults were willing to participate in the study, the appointments for home visit (to collect pictures and stories) were in some cases cancelled or postponed because grandparents were sick, had a doctor's visit, traveled to the countryside, or were celebrating holidays. Therefore, researchers need to consider potential delays in data collection when developing the time frame of the study.

Enjoyable But Lengthy Conversation With Grandparents

Older adults were willing to share photos and stories and enjoyed the sessions. Because older adults were eager to share their stories and photos and the interviewers were attentive to this situation and did not rush the interviews, the duration of the conversations was much longer than the 1 hour initially allocated. We recommend researchers account for this aspect when scheduling reminiscence sessions.

Managing Preferences While Ensuring Diversity in the Pictures

When selecting the pictures for the study, older adults expressed preferences on certain pictures to share, which affected our diversity criteria for the pictures and stories. To address this situation, we recommend not imposing the choice on the participants but expanding the number of pictures from which to select.

Sharing Pictures on Social Media Raised No Privacy Concerns

Grandparents expressed no privacy concerns in sharing the pictures, stating that they knew the pictures were meant to be shared only with their grandchildren. We recommend being transparent about how the pictures will be shared and treated.

Managing the Picture and Story Sharing Proved Feasible on Facebook

The setup consisted of a closed group on Facebook accessible only to the grandchildren. This setting was accepted by the participants, and the sharing was carried out without any difficulties. Pictures were shared every week, and grandchildren were interactive using Facebook reactions such as thumbs up or hearts as feedback on the posts. A few of the participants also expressed their gratitude by commenting on the posts.

Running the Proposed Study Process Is Feasible

The grandchildren could easily understand and complete the questionnaires. There were no complaints regarding the questionnaire or the set up of the Facebook groups. In addition, some of the grandchildren said that they enjoyed being part of the study because it reminded them about positive memories of their childhood as well as their grandparents. There were no dropouts.

Alternative Study Designs

We identified two potential issues that led us to propose alternative study designs. The first issue was balancing the intervention and control group (equipoise). The grandchildren who agree to participate in this study will likely do so to experience the intervention. Thus, assigning them to the control group would not fulfill their decision to be involved in this study. To address this potential issue, we propose a crossover design where control and intervention groups would be swapped after the 2-month intervention period so that the control group becomes the intervention group and vice versa. After the crossover, the same intervention process would be applied for another 2 months, allowing all participants to experience the intervention. Incidentally, the second intervention period would serve as a follow-up for the intervention group. The second issue was low rate of recruitment. In case the researchers are unable to randomize 60 pairs of participants at the same time due to low rate of recruitment, a quasiexperimental design would provide flexibility in choosing the control group, allowing those who agree to participate in the study to experience the intervention (to join the intervention group). The potential selection bias in this alternative should be noted, however. May any of the above changes become necessary, amendments to the current version of the protocol will be applied and communicated to the relevant venue.

Discussion

Summary

This protocol presents the design of an RCT to assess effects of sharing old pictures with grandchildren on intergenerational relationships with further recommendations to execute the protocol in future studies.

Strengths

Maintaining intergenerational relationship over time is beneficial to both young and older adults. However, very few interventions focus on the specific challenges of intergenerational interactions, especially from the young adult side. This study proposes an intervention to maintain and improve the relational quality in the grandparent-grandchild relationship using social media. Social media are widely used among the youth and are therefore a promising resource to discover interesting aspects of the life of their loved ones, which could potentially translate in more conversational resources and motivate an increased commitment in the relationship. Our proposed intervention can be potentially

implemented to any society with the technological means, without the need for sophisticated tools.

Limitations

There are number of possible limitations to this study. First, unsuccessful recruitment of study participants, which led us to reconsider the study design, can be a limiting factor. However, the issue can be attributed to the specific demographical characteristics of Mongolia, as the most sparsely populated country, and other countries may see better chances at successful recruitments. In addition, we proposed an RCT design to identify the impact of social media in the intergenerational relationship instead of a more robust multiple armed RCT design. Although the latter would have been more effective in identifying other factors influencing the intervention, its application would have required a larger population sample, which would have added further challenges to an already complicated recruitment process. Therefore, we opted for collecting qualitative data at the end of the study (debriefing session) to identify additional factors in the intervention, thus mitigating the risks involved in recruiting a larger sample. Furthermore, the study will be carried out only in the city; therefore, people who live in rural areas are not included. Finally, the recruitment will be implemented by reaching out only to the grandchildren; therefore, the future studies will consider involving grandparents in the recruitment in order to reduce the selection bias.

Conclusion

This paper presented the design of an RCT to assess the effects of sharing old pictures with grandchildren on intergenerational relationships. The intervention leverages the practice of reminiscence and social media sharing to empower grandchildren through the sharing of family stories. We focused on grandchildren, and particularly young adults, as changing their communication behavior and investment can have great effects on the relationship with their grandparents. The proposed intervention can be potentially implemented in any society with the technological means, without the need for sophisticated tools. The results from the pretest pilot highlight some practical recommendations for deploying the protocol.

Acknowledgments

This work has received funding from the EU Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 690962.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Project evaluation by the funding agency (European Commission). The project evaluation is based on the activities planned, including user studies.
[PDF File (Adobe PDF File), 88 KB - resprot_v9i4e16315_app1.pdf]

References

1. Kalliopuska M. Relations of retired people and their grandchildren. Psychol Rep 2016 Dec 06;75(3):1083-1088. [doi: 10.2466/pr0.1994.75.3.1083]
2. Mansson DH. Grandparents' expressed affection for their grandchildren: examining the grandparents' own psychological health. Comm Res Rep 2014 Oct 21;31(4):329-338. [doi: 10.1080/08824096.2014.963218]
3. Williams A, Giles H. Intergenerational Conversations Young Adults' Retrospective Accounts. Human Comm Res 1996 Dec;23(2):220-250. [doi: 10.1111/j.1468-2958.1996.tb00393.x]
4. Soliz J. Intergenerational Support and the Role of Grandparents in Post-Divorce Families: Retrospective Accounts of Young Adult Grandchildren. Qualitative Research Reports in Communication 2008 Oct 20;9(1):72-80. [doi: 10.1080/17459430802400373]
5. Nussbaum F, Coupland J. Handbook of Communication and Aging Research. London: Routledge; 2004.
6. Bangerter LR, Waldron VR. Turning points in long distance grandparent-grandchild relationships. J Aging Stud 2014 Apr;29:88-97. [doi: 10.1016/j.jaging.2014.01.004] [Medline: 24655676]
7. Harwood J. Communication media use in the grandparent-grandchild relationship. J Comm 2000;50(4):56-78. [doi: 10.1111/j.1460-2466.2000.tb02863.x]
8. Holladay SJ, Seipke HL. Communication between grandparents and grandchildren in geographically separated relationships. Comm Stud 2007 Sep 05;58(3):281-297. [doi: 10.1080/10510970701518371]
9. Williams A, Nussbaum F. Intergenerational Communication Across the Life Span. London: Routledge; 2013.
10. Harwood J. Communicative predictors of solidarity in the grandparent-grandchild relationship. J Soc Personal Relat 2016 Jun 30;17(6):743-766. [doi: 10.1177/0265407500176003]
11. Fowler C, Soliz J. Responses of young adult grandchildren to grandparents' painful self-disclosures. J Lang Soc Psychol 2009 Nov 18;29(1):75-100. [doi: 10.1177/0261927x09351680]

12. Lin M, Harwood J, Bonnesen JL. Conversation topics and communication satisfaction in grandparent-grandchild relationships. J Lang Soc Psychol 2016 Jul 26;21(3):302-323. [doi: 10.1177/0261927x02021003005]
13. Mills TL. When grandchildren grow up: role transition and family solidarity among baby boomer grandchildren and their grandparents. J Aging Stud 1999 Jun;13(2):219-239. [doi: 10.1016/s0890-4065(99)80052-8]
14. Forghani A, Neustaedter C. The routines and needs of grandparents and parents for grandparent-grandchild conversations over distance. Proc of the SIGCHI Conference on Human Factors in Computing Systems 2014. [doi: 10.1145/2556288.2557255]
15. Strom RD, Strom PS. Assessment of intergenerational communication and relationships. Educat Gerontol 2014 May 21;41(1):41-52. [doi: 10.1080/03601277.2014.912454]
16. Choi M, Kong S, Jung D. Computer and internet interventions for loneliness and depression in older adults: a meta-analysis. Healthc Inform Res 2012 Sep;18(3):191-198 [FREE Full text] [doi: 10.4258/hir.2012.18.3.191] [Medline: 23115742]
17. Chen YR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18 [FREE Full text] [doi: 10.2196/jmir.4596] [Medline: 26822073]
18. Morris ME, Adair B, Ozanne E, Kurowski W, Miller KJ, Pearce AJ, et al. Smart technologies to enhance social connectedness in older people who live at home. Australas J Ageing 2014 Sep;33(3):142-152. [doi: 10.1111/ajag.12154] [Medline: 24730370]
19. Welsh D, Morrissey K, Foley S, McNaney R, Salis C, McCarthy J, et al. Ticket to talk: supporting conversation between young people and people with dementia through digital media. Proc of the 2018 CHI Conference on Human Factors in Computing Systems 2018. [doi: 10.1145/3173574.3173949]
20. Cornejo R, Tentori M, Favela J. Enriching in-person encounters through social media: a study on family connectedness for the elderly. Int J Hum Comput Stud 2013 Sep;71(9):889-899. [doi: 10.1016/j.ijhcs.2013.04.001]
21. Laconich J, Tranquillini S, Casati F, Nikolaevna K, Nervez K. Improving family connectedness through the sharing of life experiences. 2014. URL: https://stefanotranquillini.com/papers/JaraAVI2014.pdf [accessed 2020-03-12]
22. Brewer RN, Jones J. Pinteresce: exploring reminiscence as an incentive to digital reciprocity for older adults. Proc of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing 2015. [doi: 10.1145/2685553.2699017]
23. Strom R, Strom P, editors. Grandparents and reciprocal learning for family harmony. In: Achieving Quality Education for All. Berlin: Springer; 2013:139-145.
24. Ibarra F, Baez M, Fiore F, Casati F. Designing for co-located and virtual social interactions in residential care. In: Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems. New York, NY, USA: Association for Computing Machinery; 2018 Presented at: Designing Interactive Systems Conference; 2018; Hong Kong p. 129-134. [doi: 10.1145/3197391.3205424]
25. Ibarra F, Baez M, Fiore F, Casati F. Stimulating conversations in residential care through technology-mediated reminiscence. 2017 Presented at: INTERACT 2017: 16th IFIP TC 13 International Conference; 2017; Mumbai p. 62-71. [doi: 10.1007/978-3-319-67687-6_5]
26. Webster JD, Gould O. Reminiscence and vivid personal memories across adulthood. Int J Aging Hum Dev 2007;64(2):149-170. [doi: 10.2190/Q8V4-X5H0-6457-5442] [Medline: 17451043]
27. Webster JD. Construction and validation of the Reminiscence Functions Scale. J Gerontol 1993 Sep;48(5):P256-P262. [doi: 10.1093/geronj/48.5.p256] [Medline: 8366271]
28. Westerhof GJ, Bohlmeijer ET. Celebrating fifty years of research and applications in reminiscence and life review: state of the art and new directions. J Aging Stud 2014 Apr;29:107-114. [doi: 10.1016/j.jaging.2014.02.003] [Medline: 24655678]
29. Kemp S. Digital 2019: Mongolia. 2019 Jan 31. URL: https://datareportal.com/reports/digital-2019-mongolia [accessed 2020-03-12]
30. Gächter S, Starmer C, Tufano F. Measuring the closeness of relationships: a comprehensive evaluation of the "inclusion of the other in the self" scale. PLoS One 2015;10(6):e0129478 [FREE Full text] [doi: 10.1371/journal.pone.0129478] [Medline: 26068873]
31. Buhrmester D, Furman W. The Network of Relationships Inventory: Relationship Qualities Version. 2008. URL: https://www.midss.org/sites/default/files/network_of_relationships_questionnaire_manual.6.21.2010_1.doc [accessed 2020-03-12]
32. Rusbult CE, Martz JM, Agnew CR. The Investment Model Scale: measuring commitment level, satisfaction level, quality of alternatives, and investment size. Personal Relationships 1998 Dec;5(4):357-387. [doi: 10.1111/j.1475-6811.1998.tb00177.x]
33. Campbell K, Holderness N, Riggs M. Friendship chemistry: an examination of underlying factors. Soc Sci J 2019 Dec 09;52(2):239-247. [doi: 10.1016/j.soscij.2015.01.005]

Abbreviations

ANOVA: analysis of variance
FCQ: Friendship Chemistry Questionnaire

IMS: Investment Model Scale
IOS: Inclusion of Other in the Self scale
NRI-RQV: Network of Relationships Inventory–Relationship Quality Version
RCT: randomized controlled trial
RQ1: research question 1
RQ2: research question 2
RQ3: research question 3

Edited by G Eysenbach; submitted 18.09.19; peer-reviewed by MA Ferrario, S Fornasini; comments to author 28.10.19; revised version received 07.11.19; accepted 14.01.20; published 30.04.20.

Please cite as:
Gantumur Z, Baez M, Ulamnemekh NE, Ibarra F, Myagmarjav S, Casati F
Effects of Sharing Old Pictures With Grandchildren on Intergenerational Relationships: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e16315
URL: http://www.researchprotocols.org/2020/4/e16315/
doi:10.2196/16315
PMID:32352393

Original Paper

A Smartphone App (mDASHNa-CC) to Support Healthy Diet and Hypertension Control for Chinese Canadian Seniors: Protocol for Design, Usability and Feasibility Testing

Ping Zou[1], PhD; Jennifer Stinson[2], PhD; Monica Parry[3], PhD; Cindy-Lee Dennis[4], PhD; Yeqin Yang[5], PhD; Zhongqiu Lu[5], PhD

[1]School of Nursing, Nipissing University, Toronto, ON, Canada

[2]Lawrence Bloomberg Faculty of Nursing, University of Toronto, Hospital for Sick Children, Toronto, ON, Canada

[3]Lawrence Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

[4]Lawrence Bloomberg Faculty of Nursing and Department of Psychiatry, University of Toronto, Toronto, ON, Canada

[5]School of Nursing, Wenzhou Medical University, Wenzhou, China

Corresponding Author:
Ping Zou, PhD
School of Nursing
Nipissing University
750 Dundas Street West, Room 209
Toronto, ON, M6J 3S3
Canada
Phone: 1 416 642 7003
Email: pingz@nipissingu.ca

Abstract

Background: This proposed study aims to translate the Dietary Approach to Stop Hypertension with Sodium (Na) Reduction for Chinese Canadians (DASHNa-CC), a classroom-based, antihypertensive, dietary educational intervention, to an innovative smartphone app (mDASHNa-CC). This study will enable Chinese Canadian seniors to access antihypertensive dietary interventions anytime, regardless of where they are. It is hypothesized that senior Chinese Canadians will be satisfied with their experiences using the mDASHNa-CC app and that the use of this app could lead to a decrease in their blood pressure and improvement in their health-related quality of life.

Objective: The goal of this study is to design and test the usability and feasibility of a smartphone-based dietary educational app to support a healthy diet and hypertension control for Chinese Canadian seniors.

Methods: A mixed-method two-phase design will be used. The study will be conducted in a Chinese immigrant community in Toronto, Ontario, Canada. Chinese Canadian seniors, who are at least 65 years old, self-identified as Chinese, living in Canada, and with elevated blood pressure, will be recruited. In Phase I, we will design and test the usability of the app using a user-centered approach. In Phase II, we will test the feasibility of the app, including implementation (primary outcomes of accrual and attrition rates, technical issues, acceptability of the app, and adherence to the intervention) and preliminary effectiveness (secondary outcomes of systolic and diastolic blood pressure, weight, waist circumference, health-related quality of life, and health service utilization), using a pilot, two-group, randomized controlled trial with a sample size of 60 participants in a Chinese Canadian community.

Results: The study is supported by the Startup Research Grant from Nipissing University, Canada. The research ethics application is under review by a university research ethics review board.

Conclusions: The study results will make several contributions to the existing literature, including illustrating the rigorous design and testing of smartphone app technology for hypertension self-management in the community, exploring an approach to incorporating traditional medicine into chronic illness management in minority communities and promoting equal access to current technology among minority immigrant senior groups.

Trial Registration: Clinicaltrials.gov NCT03988894; https://clinicaltrials.gov/ct2/show/NCT03988894

International Registered Report Identifier (IRRID): PRR1-10.2196/15545

(JMIR Res Protoc 2020;9(4):e15545) doi:10.2196/15545

KEYWORDS

diet; hypertension; smartphone app; senior; Chinese; Canada

Introduction

Rationale for the Intervention

In Canada, 1.3 million Chinese individuals comprise approximately 4.0% of Canada's population and 21.2% of the country's visible minorities [1]. In this Chinese population, hypertension is the most prominent risk factor for cardiovascular disease and accounts for a large proportion of stroke [2] and heart failure [3]. With a 15.1% hypertension prevalence rate [4], Chinese Canadians are at an increased risk of cardiovascular disease and associated morbidity and mortality. Because the prevalence of hypertension increases with age, Chinese Canadian seniors are at especially high risk for hypertension and related mortality [4].

Compared with antihypertensive dietary suggestions in hypertension care guidelines [5], Chinese Canadians have a suboptimal dietary intake, which impacts their blood pressure control. Chinese Canadians' sodium intake is higher than antihypertensive dietary recommendations [6], and a high proportion of Chinese Canadians consume fewer fruits and vegetables than antihypertensive dietary recommendations [7,8]. In addition, Chinese Canadians have a low dairy intake compared with antihypertensive dietary recommendations [9]. Dietary factors have been identified as the most important risk factors for hypertension among the Chinese population [10]. As such, effective dietary interventions are needed to assist with blood pressure control in Chinese Canadian seniors [11,12].

Dietary Interventions for Hypertension Control

The Dietary Approach to Stop Hypertension (DASH) and sodium reduction are antihypertensive dietary interventions recommended by Canadian hypertension care guidelines [5,13]. Focused on healthy dietary patterns rather than specific nutrients, the DASH diet includes eight food groups (grains, vegetables, fruits, meats, dairy, nuts, fats, and sweets), with specific serving suggestions [14]. The DASH studies demonstrate an effective systolic blood pressure reduction of 4-11 mmHg and a diastolic blood pressure reduction of 2-6 mmHg [15-19]. A systematic review has suggested that sodium reduction is also related to decreased blood pressure [20]. However, the DASH and sodium reduction interventions do not consider the psychosocial factors that influence dietary behaviors, nor have they been tested among Chinese Canadians in a community setting [5,15]. A review paper explored the cultural factors influencing diet and hypertension control in the Chinese Canadian population and suggested that English language proficiency, health literacy, traditional Chinese diet, migration and acculturation, and traditional Chinese medicine (TCM) influence Chinese Canadians' dietary practices [21]. A culturally tailored intervention is thus needed to facilitate blood pressure control in Chinese Canadians.

Incorporation of Traditional Chinese Medicine for Hypertension Control

Chinese Canadians rely strongly on TCM for chronic illness management and prefer to incorporate TCM into their health care [11]. Food therapy is an essential component of TCM and has been acknowledged as a successful therapy for more than 3000 years [22]. In TCM, food is conceptualized with both nutritional and functional considerations. Like medicines, food can be used to maintain health, prevent and treat diseases, and facilitate rehabilitation [23-27]. There are four principles of TCM food therapy, including light eating, balancing the hot and cold nature of food, harmony of the five flavors of food, and consistency of diets with different health conditions [28,29]. The principal investigator of this study has previously published a literature review on TCM food therapy and hypertension control using a rigorous review method and statistical analysis [30]. Findings suggest that some foods have antihypertensive functions [31-41], and food therapy can facilitate hypertension control [42-45].

Antihypertensive Diet Apps for the Chinese Population

We conducted a scoping review of existing antihypertensive diet apps written in Chinese on the current market. We searched in various app stores and found 24 apps (15 on iTunes, seven on the Google Play Store, and two on the Chinese App Market). All apps were written in Chinese and focused on diets for hypertension control. We screened the app description and conducted a content analysis. We identified several gaps, including: (1) the app producers had no credentials from any licensed health care professional or research team; (2) all apps were in electronic book form, only providing information with no user interactivity; (3) none of the apps focused on senior users; and (4) most apps from iTunes and the Google Play Store were based on Western diets, which may not fit with Chinese diets. The findings from our review are consistent with a review of app studies on hypertension control, which stated that most of the current apps lack standardization and scientific validation [46]. The mobile Dietary Approach to Stop Hypertension with Sodium (Na) Reduction for Chinese Canadians (mDASHNa-CC) app differs from other related apps in the current market in that it will be developed by a team of health care professionals, involve community end-users in the development process, be based on the current gold standard of antihypertensive dietary interventions, incorporate TCM to ensure that it is culturally significant for Chinese seniors, provide immediate response and recommendations according to the patients' current conditions, and be scientifically tested by a Randomized Control Trial.

Preliminary Work: Success of the DASHNa-CC Pilot Trial

The Dietary Approach to Stop Hypertension with Sodium Reduction for Chinese Canadians (DASHNa-CC) intervention was designed based on current literature and clinical expertise.

The DASHNa-CC integrates TCM food therapy into the DASH and sodium reduction diary intervention for blood pressure control. Adapted from DASH, DASHNa-CC is designed as a standardized, culturally sensitive, dietary education intervention for Chinese Canadians. The contents of the DASHNa-CC intervention have three components: (1) the DASH diet pattern, including characteristics of the DASH diet, serving size estimation tool, foods rich in calcium, and foods rich in potassium; (2) sodium reduction, including the importance of sodium reduction for cardiac health, goals of sodium reduction, and 20 sodium reduction strategies; (3) TCM food therapy, including the contribution of TCM to hypertension control, four principles of TCM food therapy, and 34 foods and five herbal teas with antihypertensive functions recommended by TCM food therapy. The intervention delivery consisted of: (1) the DASHNa-CC Intervention Manual; (2) two classroom sessions (2 hours per session); and (3) one 20-minute booster telephone call.

From August to December 2014, 60 participants were recruited in a pilot randomized controlled trial to examine the feasibility and potential effects of the DASHNa-CC intervention. The research findings suggested that participants adhered well to the DASH diet pattern, sodium reduction, and TCM food therapy strategies. The loss to follow-up rate was 5%. Participants were highly satisfied with the intervention and perceived that the intervention contents were helpful, the delivery approaches were suitable, and their participation in the pilot trial was beneficial rather than a burden on their lives. Compared to the control group, the intervention group lowered their systolic blood pressure by 3.8 mmHg (t_{55}=–1.58; *P*=.12) more than the control and lowered their diastolic blood pressure by 2.4 mmHg (t_{55}=–1.22; *P*=.23). These blood pressure reductions were clinically important to reduce hypertension-related mortality and morbidity [47,48]. In addition, the intervention group had a significant improvement from baseline to eight weeks post-randomization in the physical component score (t_{55}=2.13; *P*=.04) of the Medical Outcomes Study 36-Item Short-Form version two (SF-36v2). It is concluded that the DASHNa-CC intervention has the potential to decrease systolic and diastolic blood pressure and improve the health-related quality of life for Chinese Canadians. Three papers, which discuss the main outcome of the DASHNa-CC study [49], the recruitment process [50], and participant satisfaction [51] to the DASHNa-CC intervention, respectively, were published.

Why is the mDASHNa-CC Smartphone App Needed?

The smartphone app version of the DASHNa-CC intervention (mDASHNa-CC) is proposed because many Chinese Canadian seniors were unable to participate in the pilot study due to busy life schedules (eg, taking care of grandchildren) and were unable to travel to the community center where the DASHNa-CC intervention was delivered. Also, in this pilot study, it was found that most Chinese Canadian seniors were well-educated, technologically savvy, able to access the internet, owned a smartphone, and were willing to learn new skills and knowledge to improve their hypertension control. The academic committee of the DASHNa-CC pilot study suggested transferring the DASHNa-CC to a home-based intervention using a website or smartphone technology to better meet the health care needs of Chinese Canadian seniors. We conducted two focus group discussions with 20 pilot study participants in the community in 2016. All participants had a smartphone and expressed an eagerness to use our app when it becomes available. They stated that they "review the manual frequently, and a smartphone app will make the manual easier to use," and "this is something new and I want to try."

Specific Aims

The overall aim of this two-phase study is to translate the Dietary Approach to Stop Hypertension with Sodium Reduction for Chinese Canadians, a classroom-based antihypertensive dietary educational intervention, to an innovative smartphone app (mDASHNa-CC). This smartphone app would enable and empower Chinese Canadian seniors' to access this antihypertensive intervention anytime, regardless of where they are.

Research Objectives and Research Questions

The research objectives are to design and test the usability and feasibility of a smartphone-based dietary educational app to support a healthy diet and hypertension control for Chinese Canadian seniors. In the Phase I usability testing study, the research questions are: (1) How can the mDASHNa-CC app be designed according to the DASHNa-CC intervention and the current literature on hypertension webpages and apps; and (2) How can the mDASHNa-CC app be refined using a user-centered design approach to ensure it is easy to use, efficient, and satisfying for participants? The Phase II pilot feasibility testing study will examine the implementation and preliminary effectiveness of using the mDASHNa-CC app with Chinese Canadians seniors who have hypertension in the community. The research questions of implementation are: (1) What are the rates of participant accrual and attrition; (2) What technical issues arise over the study; (3) What is the acceptability of the study protocol; (4) To what extent do participants adhere to the requirements of learning dietary educational material, conducting dietary self-assessments, and monitoring blood pressure with the mDASHNa-CC app; and (5) To what extent do participants adhere to the dietary recommendations in the mDASHNa-CC app? The research questions of preliminary effectiveness are: Compared to usual care, what is the effect of an 8-week mDASHNa-CC app intervention on systolic and diastolic blood pressure, body weight, waist circumference, health-related quality of life, and health service utilization?

Theoretical Framework

Social cognitive theory, a behavioral model commonly applied to the design of interventions intended for the self-directed management of chronic disease, will be used as the governing behavior change theory for the mDASHNa-CC intervention. The key constructs of this theory include psychological determinants of behavior, observational learning, environmental determinants of behavior, self-regulation, and moral disengagement [52]. These constructs emphasize the dynamic interaction of personal, behavioral, and environmental factors that could alter human behavior [52]. Social cognitive theory

has been successfully used to guide the app design for diabetes control in Canada [53]. Learning from this Canadian evidence, we applied social cognitive theory to our app design for antihypertensive dietary behavior self-management. Evidence suggests that psychological and environmental determinants impact Chinese adults' dietary behavior and hypertension control. We hypothesized that promoting positive psychological and environmental determinants can enhance dietary behavior and blood pressure control. Learning is considered as the foundational function for the mDASHNa-CC app; therefore, antihypertensive dietary education acts as an essential component of the app. The core function of the app is self-regulation; thus, dietary self-assessments and blood pressure self-monitoring are embedded in this app as the main functions. In addition, automatic feedback according to dietary self-assessments and blood pressure data is incorporated in the app to encourage health behavior changes. This app design will drive self-efficacy through observational learning, self-regulation, and incentive motivation. The behavior changes will be achieved through self-monitoring, tailored feedback, structured education, and incentivizing positive behavior.

Methods

Phase I: Design and Usability Testing of the mDASHNa-CC App

Adhering to a phased sequential approach to the development of complex technology-based interventions [54], the development of the mobile Dietary Approach to Stop Hypertension with Sodium (Na) Reduction for Chinese Canadians (mDASHNa-CC) app will include app design, usability testing, and feasibility testing. A user-centered design approach will be applied, where seniors will be actively engaged in all aspects of the research process, including the app's design, the usability and feasibility testing, and the refinement of the prototype.

Convenience sampling will be used to recruit eligible individuals. To access the most representative sample of Chinese Canadian seniors, a community center in Toronto where Chinese Canadians occupy a high percentage of the total population will be used for blood pressure screening, participant recruitment, app testing, and participant follow-up. In partnership with the Ontario Chinese Senior Association, we will host blood pressure screening events in the Chinese community to facilitate the recruitment process. The app testing procedures will commence after participation eligibility is assessed, informed consent is obtained, and a trained research assistant collects demographic and other outcome data.

This study will include all Chinese Canadians who: (1) are at least 65 years old; (2) have a systolic blood pressure higher than 140 mmHg, or a diastolic blood pressure higher than 90 mmHg, or are on antihypertensive medications, based on preintervention baseline assessment; (3) can understand (listen) and speak in Mandarin or Cantonese, and can read and write in Chinese; and (4) have access to a smartphone. Since self-reporting has been recommended as the preferred approach to measure ethnicity and has been widely applied in public health studies [55,56], the identification of Chinese Canadians in this study will be based on self-reporting of ethnicity. The study will exclude individuals who: (1) have special dietary requirements; (2) are a household member of another mDASHNa-CC participant; or (3) plan to leave the area before the anticipated end of the study.

App Design

Based on the findings of the DASHNa-CC pilot trial, the major functions of the app will include: (1) antihypertensive dietary education; (2) dietary self-assessments; (3) automatic feedback according to dietary self-assessments; (4) blood pressure monitoring; and (5) automatic feedback according to blood pressure data. To enhance individuals' interest in using the app, a new function of the app that will be utilized is built-in age and culture-specific entertainment content as a reward for learning dietary education material, conducting dietary self-assessment, and monitoring blood pressure [57]. The entertainment content, including Chinese songs, videos, or Beijing opera, will be suggested by senior Chinese Canadians in the community.

The antihypertensive dietary education is based on the DASHNa-CC intervention and includes the content of the DASH diet pattern, sodium reduction, and TCM food therapy. It is recommended that participants review the educational content and answer the related questions, and they will also be asked to conduct a dietary self-assessment every day on the app. Based on this, the app will automatically provide feedback and suggestions. Participants are also asked to measure their blood pressure twice a day using a home blood pressure monitor. They will then record the data in the app, which will automatically provide feedback regarding their blood pressure status. Frequent use of the app will be rewarded with built-in entertainment content. The data entered by seniors will be stored locally on the smartphone and then communicated to the server when the phone is online using an encrypted protocol. The server will be hosted at Nipissing University behind a firewall in a secure network environment. A username and password will be required to access the data. The app development will be completed by a software programmer who has extensive experience in educational smartphone app development in the Chinese Canadian community.

Usability Testing

Usability testing is a widely used methodology that incorporates an iterative process of testing an intervention's user-interface and then applying the results to redesign the prototype to meet users' needs. The current literature supports the importance of usability testing to increase the likelihood of a technology-based intervention's effectiveness [58]. It is recommended that usability testing takes two to three cycles and involves five to seven participants in every cycle [57,59].

Low Fidelity User-Centered Design

A qualitative usability testing approach will be used with multiple iterative cycles of semistructured audiotaped interviews. The app interface designs will be trialed with participants. A total of seven seniors in each cycle will be shown paper screenshots of the application and asked what they like and dislike about the interface design, contents, major functions, and built-in entertainment. The list of interview questions will

be modified during the interview process considering emerging themes and field notes related to perceived ease of use. The research assistant will record problems with the app. Seniors will also be asked to provide further suggestions for improvement. Design elements will be modified, and new paper screenshots will be generated and tested with iterative cycles until no further changes are suggested.

High Fidelity User-Centered Design

Following the development of a fully functional smartphone-based prototype, usability testing (multiple iterative cycles) with semistructured audiotaped interviews will be conducted again with seven seniors in each cycle. In this phase, a trained research assistant will first provide seniors with a brief (approximately 5 minutes) demonstration of the app on a smartphone using a standardized dietary assessment vignette. Seniors will then be asked to complete a dietary self-assessment in the app and record their food intake while thinking aloud about their likes, dislikes, and difficulties with the app. At the end of each session, a research assistant will ask a series of open-ended questions related to ease of use, what seniors liked or disliked about the app, and any technical issues. The research assistant will record the answers to questions, write field notes on ease of app use, and explore emerging themes. After the first iterative cycle, changes will be made based on the themes identified from seniors' opinions. Conflicting suggestions will be handled based on the majority. Another iterative cycle will be conducted with another seven seniors until there are no further recommendations for change to the app.

Measurement Tools

The baseline participant demographic characteristics will be collected via the Participant Information Questionnaire. This questionnaire includes 23 questions about socioeconomic status, risk factors for hypertension, and migration history. This questionnaire was used in the DASHNa-CC pilot study. In addition, overall comfort level with smartphones will be ascertained using a questionnaire about smartphone ownership, level of use, and likeability. This questionnaire has been used successfully in previous app studies [57].

Data Analysis

Demographic data will be analyzed using the software SPSS version 20.0 (IBM Corporation, New York, United States). To describe the sample, various descriptive statistics (eg, means, standard deviations, proportions), dependent on the level of measurement of the variables, will be calculated for sample demographics and other baseline information.

In both low and high-fidelity usability testing, audiotaped usability interviews will be transcribed verbatim. All transcripts from the usability testing phases will be verified against the tapes and imported into the software NVivo 10.0 (QRS International, Chatstone, Australia) for coding. The field notes taken during the interviews will also be transcribed and included in the analytic process. By using thematic coding, data will be coded according to the study objectives and categorized to reflect the emerging themes [60]. Any changes to the prototype will be made based on feedback from each iterative cycle of testing.

Phase II: Feasibility Testing of mDASHNa-CC App

Following usability testing, a pilot randomized controlled trial [61] feasibility study will be conducted with Chinese Canadian seniors to determine implementation (primary outcomes, including accrual and attrition rates, technical issues, acceptability of the app, and adherence to the intervention,) and preliminary effectiveness (secondary outcomes, including systolic and diastolic blood pressure, weight, waist circumference, health-related quality of life, and health service utilization). This study is designed as a pilot two-group (1:1) randomized controlled trial with a sample size of 60 participants (block of 20) in a Chinese Canadian community in the Greater Toronto Area.

The sampling procedures and setting will be the same as the Phase I study. A convenience sample of 60 Chinese Canadian seniors will participate in this study. Self-identified Chinese Canadians were recruited if they met the following inclusion criteria: (1) at least 65 years of age; (2) had a systolic blood pressure between 140 to 159 mmHg or a diastolic blood pressure between 90 to 99 mmHg; (3) were able to understand and speak Mandarin and read and write Chinese; and (4) had access to a smartphone. Individuals were excluded if they: (1) used antihypertensive medications, insulin, or oral hypoglycemic agents; (2) had a cardiovascular event during the previous three months; (3) had a history of congestive heart failure; (4) had a cancer diagnosis or had undergone cancer treatment during the past two years; or (5) had special dietary requirements.

As a pilot study is not powered to be a hypothesis-testing trial, formal sample size calculations are not recommended [62]. Instead, the sample size suggested was based on recommendations for feasibility trials [63]. In this pilot study, 60 eligible participants will be recruited. Following university ethics approval, participants will be recruited by blood pressure screening events in diverse community settings. The recruitment process of Phase II is the same as that of Phase I. Study procedures will commence after eligibility is assessed, informed consent is obtained, and demographic and baseline outcome data are collected by a trained research assistant.

Using an online randomization tool provided by Interrand Company, Ottawa, Canada, participants will be randomized into either an intervention group or a control group. Participants randomized to the intervention group will receive the mDASHNa-CC app intervention for eight weeks plus usual care; participants randomized to the control group will receive usual care. Coinvestigators, collaborators, and outcome assessors will be blinded to the group assignment. Because this study is an app educational intervention, it will be impossible to blind participants to the group assignment. The group assignment will be concealed until all outcome data are collected [64]. After all outcome data are collected eight weeks post-randomization, participants in the control group will also be offered use of the app.

The control group will be usual care. Usual care consists of three parts: (1) receiving a general hypertension health education booklet from the Heart and Stroke Foundation of Ontario; (2) being encouraged to see their family physicians or primary health care providers regarding their blood pressure status (those

who do not have a primary health care provider will be referred to a walk-in clinic or a community health center); and (3) having access to family physicians, telehealth, emergency care, hospitals, and other health care facilities in the Greater Toronto Area as required.

In addition to usual care, those participants randomized to the intervention group will be offered use of the app. A trained research assistant will teach them how to use it, they will load the app on their smartphones, and then they will be requested to review educational material, conduct dietary self-assessments, and monitor blood pressure for eight weeks. Telephone assistance from a trained research assistant will be available to seniors in case of technical problems or if any questions arise about the app. The research team will conduct a daily review of a summary of each senior's report so that participant safety issues can be identified and resolved. By the end of the eight weeks postrandomization, seniors will be prompted by phone using an audible alert to complete the app evaluation questionnaire on their smartphone, which will ascertain likes and dislikes with the app.

Measurement Tools

Baseline demographic characteristics will be collected via the participant information questionnaire, which is described in the measurement tools section in the Phase I study. Implementation outcomes will describe the feasibility of using the app with Chinese Canadian seniors in the community. Implementation will be measured as:

- Accrual and attrition rates: The mDASHNa-CC recruitment log has been designed to record data related to the number of eligible seniors per recruitment day, reasons for ineligibility, and reasons for nonparticipation. The mDASHNa-CC Activity Log has been designed to record data on attrition, including occurrence/reasons for attrition, technical difficulties, adherence, and outcome measure completion. A trained research assistant will complete the logs daily during the research process.
- Technical issues: The occurrence and description of technical problems will be recorded on the mDASHNa-CC activity log by a trained research assistant.
- Acceptability: The acceptability e-scale ascertains perceptions related to how helpful, difficult, and enjoyable electronic-based programs are to use, how understandable questions are, and how acceptable the time invested in reporting was [65]. This scale demonstrated validity and reliability in various prior studies [65]. For the present study, the wording of the scale will be slightly modified, and a free-text question, where seniors are encouraged to enter any other information that they feel would be important to discuss, will be added. Seniors in the intervention group will fill out this scale by email four weeks and eight weeks postrandomization.
- Adherence: A built-in number counter will measure participants' adherence to the requirements of learning dietary educational material, conducting dietary self-assessments, and monitoring blood pressure in the mDASHNa-CC app. The number counter will record the time and frequency of app use. The dietary intake will be measured at baseline and eight weeks postrandomization by a one-to-one dietary interview using the validated Automated Multiple Pass Method [66] by a trained research assistant in the community center. This approach was successfully tested in the DASHNa-CC pilot trial. Adherence to the DASH will be measured by the validated DASH component score of each food group, and the total DASH score [67]. Adherence to the sodium reduction will be measured by a 24-hour urine test [68]. Adherence to TCM food therapy will be measured using a 24-statement questionnaire on a 5-point Likert scale, which was validated in the DASHNa-CC pilot study [69].

Preliminary Effectiveness and Outcomes

Except for health service utilization, which will be measured only at eight weeks postrandomization, the other following outcomes will be measured at baseline and eight weeks postrandomization at the community center by a trained research assistant during a one-to-one appointment. If a participant cannot visit the center, a home visit by a trained research assistant will be arranged.

- Systolic and diastolic blood pressures: Systolic and diastolic blood pressures will be measured with the home blood pressure monitor, Omron BP785, whose validity and reliability have been tested [70]. All blood pressure measurements will be performed following the recommended techniques by the Canadian Hypertension Education Program guidelines [5]. Each participant will be offered an Omron BP785 and training on how to measure blood pressure at home and how to record the results; however, blood pressure at baseline and eight weeks postrandomization will be measured by a trained research assistant.
- Bodyweight: An electronic body weight scale will measure weight.
- Waist circumference: Waist circumference will be measured by measurement tape following proper techniques [71-74].
- Health-related quality of life: Health-related quality of life will be measured by the SF-36v2 [75,76].
- Health service utilization: Health service utilization data will be collected via the Health Service Utilization Questionnaire, which was modified from the Health Service Utilization Questionnaire previously used in Ontario [77].

Data Analysis and Statistical Methods

Demographic data will be analyzed using the software SPSS version 20.0 by a biostatistician. To describe the sample, various descriptive statistics (eg, means, standard deviations, proportions), dependent on the level of measurement of the variables, will be calculated for sample demographics and other baseline information. Descriptive statistics will be calculated to demonstrate how participants adhered to and are satisfied with the app. Open-ended questions will be reviewed by two researchers independently. Qualitative data will be organized into meaningful groups, combining similar patterns into themes. In feasibility testing, adherence is defined as 100% when 28/28 blood pressure entries and 14/14 dietary self-assessments entries are completed within the two weeks. An independent, two-sample, two-tailed *t* test will be used to examine the

differences in mean change scores between the control and intervention groups regarding blood pressure, weight, waist circumference, and health-related quality of life.

Quality Control

Quality control strategies will be implemented. Onsite training will be provided to the research staff. The training will include research ethics, privacy and confidentiality, orientation to research protocol and procedure, literature search and review, data collection tools, data analysis methods, and community networking. The training will provide staff with adequate time for interactive learning, on-site practice, and skill preparation for comprehensive teamwork in the project [78]. To make sure there is consistency in data collection, a uniform data collection tool will be used. The principal investigator will work closely with research staff in the data collection and analysis process. Team meetings will be conducted every month for progress updates and problem-solving. Research staff will be requested to make research notes daily. Discussion and debriefing will be provided promptly if needed. To sustain participant motivation to the project, patient contact will be organized in patient-preferred time to encourage participation. A research assistant will remind participants in advance of each research event [79]. In addition, city public transport tickets will be provided to assist with commuting costs. Participation certificates will be provided to honor participants' contributions to the study.

Results

The study is supported by the Startup Research Grant from Nipissing University, Canada. The research ethics application is under review by a university research ethics review board.

Discussion

Knowledge Translation

Our knowledge translation plan incorporates strategies to ensure that our app will stand apart from existing apps in the eyes of Chinese Canadian seniors and other Chinese populations. Firstly, our networks with Chinese Canadian communities and our engagement of key consumer groups (Ontario Chinese Senior Association and other Chinese senior groups and community centers) will help to spread awareness of our app to people in the community using a grassroots approach. Collaboration with the community to provide workshops, information sessions, support groups, and social media interviews will promote the use of our app in the community. Secondly, our research team includes a nurse, a dietitian, a TCM practitioner, and a medical doctor who are in various organizations in Canada. Upon completion of the project, these leaders will be able to endorse the uptake of our app at their clinics and the practices of their colleagues. Thirdly, we will enter our app into competitions for technology design awards nationally and internationally to further solidify its credibility and earn patient buy-in. This strategy can also help disseminate our app outside Canada, including the United States, where 2.4 million Chinese individuals live, and East Asian countries, such as China, where hypertension is emerging as a critical public health issue [80,81]. Fourthly, we will partner with Hypertension Canada, Toronto Public Health, Wenzhou Medical School (P. R. China), and other organizations to promote the use of our app. Endorsement of an app by health promotion organizations can lend legitimacy to a new tool. Fifthly, national, and international academic audiences will be reached through academic publications and presentations at conferences by researchers. Finally, fact sheets, research summaries, presentations in leadership forums, and policy recommendations will be used to communicate research findings with government and policymakers to promote related policy changes.

Human Subjects

Chinese Canadian seniors will be the research subjects in this study. As a technology-based dietary educational intervention, this study poses no known risks to participants. All participants have access to telehealth, emergency care, hospitals, and other health care facilities in the Greater Toronto Area. All participants are free to use all these health care services anytime, as needed.

There are no known benefits to participation in this pilot trial. However, participants in the DASH trial reported reduced blood pressure and enhanced quality of life [15,82]. Participants in a trial of TCM food therapy in China also reported improved health-related quality of life and reduced the use of their antihypertensive medications [45]. By participating in this study, participants will gain knowledge about healthy eating and the importance of blood pressure control. Participants will be instructed to monitor their blood pressure. In addition, a CAD$20 ($15) dollar gift card and a certificate of participation will be offered to all participants as a token of appreciation for their participation. Every participant will be offered two city public transportation tickets every time they attend the research activities to compensate for travel expenses. Participants will receive the study results by email.

Implications

The study results will make contributions in six areas: (1) produce a smartphone app, which allows a large number of seniors, their families, and other community members to access the dietary intervention for hypertension control, and could potentially be used across Canada and internationally in large Chinese ethnic populations; (2) illustrate the rigorous design and testing of smartphone app technology for hypertension self-management in the community; (3) explore the approach of incorporating traditional medicine in chronic illness management in minority communities; (4) contribute to culturally sensitive care, which is an urgent need due to global migration and has implications for immigrant-recipient countries and multiethnic societies; (5) promote equal access to current technology among minority immigrant senior groups; and (6) facilitate the full randomized trial in the future to examine the effects of the app on blood pressure and health-related quality of life.

Acknowledgments

This study is supported by the Startup Research Grant, Nipissing University, Canada.

Authors' Contributions

PZ conceptualized the project and drafted the manuscript. JS, MP, CLD, YY, and ZL reviewed and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Statistics C. Immigration and Ethnocultural Diversity in Canada: National Household Survey 2011. Ottawa: Ministry of Industry; 2013. URL: https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001-eng.pdf [accessed 2020-02-18]
2. Yong H, Foody J, Linong J, Dong Z, Wang Y, Ma L, et al. A Systematic Literature Review of Risk Factors for Stroke in China. Cardiology in Review 2013;21(2):77-93. [doi: 10.1097/crd.0b013e3182748d37]
3. Moe GW, Tu J. Heart failure in the ethnic minorities. Current Opinion in Cardiology 2010;25(2):124-130. [doi: 10.1097/hco.0b013e328335fea4]
4. Chiu M, Austin PC, Manuel DG, Tu JV. Comparison of cardiovascular risk profiles among ethnic groups using population health surveys between 1996 and 2007. CMAJ 2010 May 18;182(8):E301-E310 [FREE Full text] [doi: 10.1503/cmaj.091676] [Medline: 20403888]
5. Dasgupta K, Quinn RR, Zarnke KB, Rabi DM, Ravani P, Daskalopoulou SS, Canadian Hypertension Education Program. The 2014 Canadian Hypertension Education Program recommendations for blood pressure measurement, diagnosis, assessment of risk, prevention, and treatment of hypertension. Can J Cardiol 2014 May;30(5):485-501 [FREE Full text] [doi: 10.1016/j.cjca.2014.02.002] [Medline: 24786438]
6. Zhao D, Qi Y, Zheng Z, Wang Y, Zhang X, Li H, et al. Dietary factors associated with hypertension. Nat Rev Cardiol 2011 Jul 05;8(8):456-465. [doi: 10.1038/nrcardio.2011.75] [Medline: 21727918]
7. Hislop TG, Tu S, Teh C, Li L, Low A, Taylor VM, et al. Knowledge and Behaviour Regarding Heart Disease Prevention in Chinese Canadian Immigrants. Can J Public Health 2008 May 1;99(3):232-235. [doi: 10.1007/bf03405480]
8. Taylor VM, Yasui Y, Tu S, Neuhouser ML, Li L, Woodall E, et al. Heart disease prevention among Chinese immigrants. J Community Health 2007 Oct 24;32(5):299-310. [doi: 10.1007/s10900-007-9057-5] [Medline: 17922202]
9. Lv N, Cason KL. Current Dietary Pattern and Acculturation of Chinese Americans in Pennsylvania. Topics in Clinical Nutrition 2003;18(4):291-300. [doi: 10.1097/00008486-200310000-00010]
10. Wang J, Li Y. Characteristics of hypertension in the Chinese population. Curr Hypertens Rep 2012 Oct 29;14(5):410-415. [doi: 10.1007/s11906-012-0288-1] [Medline: 22843493]
11. King KM, LeBlanc P, Carr W, Quan H. Chinese immigrants' management of their cardiovascular disease risk. West J Nurs Res 2007 Nov;29(7):804-826. [doi: 10.1177/0193945906296431] [Medline: 17526869]
12. Li W, Stewart AL, Stotts N, Froelicher ES. Cultural factors associated with antihypertensive medication adherence in Chinese immigrants. J Cardiovasc Nurs 2006;21(5):354-362. [doi: 10.1097/00005082-200609000-00005] [Medline: 16966912]
13. Registered NAOO. Nursing Management of Hypertension. 2005. URL: https://rnao.ca/bpg/guidelines/nursing-management-hypertension [accessed 2020-02-18]
14. Karanja NM, Obarzanek E, Lin P, McCullough ML, Phillips KM, Swain JF, et al. Descriptive Characteristics of the Dietary Patterns Used in the Dietary Approaches to Stop Hypertension Trial. Journal of the American Dietetic Association 1999 Aug;99(8):S19-S27. [doi: 10.1016/s0002-8223(99)00412-5]
15. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure. N Engl J Med 1997 Apr 17;336(16):1117-1124. [doi: 10.1056/nejm199704173361601]
16. Vogt TMEA, Appel LJ, Obarzanek E, Moore TJ, Vollmer WM, Svetkey LP, et al. Dietary Approaches to Stop Hypertension: rationale, design, and methods. DASH Collaborative Research Group. J Am Diet Assoc 1999 Aug;99(8 Suppl):S12-S18. [doi: 10.1016/s0002-8223(99)00411-3] [Medline: 10450289]
17. Miller ER, Erlinger TP, Young DR, Jehn M, Charleston J, Rhodes D, et al. Results of the Diet, Exercise, and Weight Loss Intervention Trial (DEW-IT). Hypertension 2002 Nov;40(5):612-618. [doi: 10.1161/01.hyp.0000037217.96002.8e] [Medline: 12411452]
18. Sacks FM, Svetkey LP, Vollmer WM, Appel LJ, Bray GA, Harsha D, DASH-Sodium Collaborative Research Group. Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. DASH-Sodium Collaborative Research Group. N Engl J Med 2001 Jan 04;344(1):3-10. [doi: 10.1056/NEJM200101043440101] [Medline: 11136953]

19. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Writing Group of the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. JAMA 2003 Apr 23;289(16):2083-2093. [doi: 10.1001/jama.289.16.2083] [Medline: 12709466]
20. He FJ, MacGregor GA. Effect of longer-term modest salt reduction on blood pressure. Cochrane Database of Systematic Reviews 2004 Jan 26(1):1-3. [doi: 10.1002/14651858.cd004937]
21. Zou P. Diet and Blood Pressure Control in Chinese Canadians: Cultural Considerations. J Immigr Minor Health 2017 Apr 17;19(2):477-483. [doi: 10.1007/s10903-016-0493-0] [Medline: 27640010]
22. Topham DL. Traditional Chinese Medicine in Orthopaedic Nursing. Orthopaedic Nursing 1999;18(6):45???52. [doi: 10.1097/00006416-199911000-00009]
23. Xu Y. Perspectives on the 21st century development of functional foods: bridging Chinese medicated diet and functional foods. Int J Food Sci Tech 2001 Mar;36(3):229-242. [doi: 10.1046/j.1365-2621.2001.t01-1-00461.x]
24. Cao Y. Introduction on History of Chinese Food Therapy (Chinese). Chinese Folk Therapy 2001;9:46-47. [doi: 10.1055/b-0034-67038]
25. Dahl M. Nutrition for Chinese populations. Health Care Food Nutr Focus 2004 Apr;21(4):8-9. [Medline: 15088473]
26. Wang B. The Yellow Emperor's Cannon Internal Medicine (Chinese). Beijing: Chinese Press of Science and Technology; 1997.
27. Kung T. Outline of the Constitutional food-adjusting (Chinese). Academy of Zhejiang Chinese Medical University 2006;30(3):217-219.
28. Zhang XL, Wu F. Regimen in Traditional Chinese Medicine (Chinese). Beijing: China Press of Traditional Chinese Medicine; 2005.
29. Liu ZC. Basic Theories of Traditional Chinese Medicine (Chinese). Beijing: High Education Press; 2007.
30. Zou P. Traditional Chinese Medicine, Food Therapy, and Hypertension Control: A Narrative Review of Chinese Literature. Am. J. Chin. Med 2016 Dec 06;44(08):1579-1594. [doi: 10.1142/s0192415x16500889]
31. Chen HZ. The Diet of Adjustment and Nutrition (Chinese version). Beijing: People Military Press; 2003.
32. Yu S. Yellow Emperor's Internal Medicine (Chinese version). Beijing: Press of Zhao Hua; 2006.
33. Xu S, Niu B. Shen Long Ben Cao Jing (Chinese Version). Shijiazhuang: The Press of Science and Technology of Hebei; 1994.
34. Li S. Ben Cao Gang Mu. Beijing: The Press of Science; 1998.
35. Zhang Z. Shang Han Zha Bing Lun (Chinese version). Shijiazhuang: The Press of Science and Technology of Hebei; 1994.
36. Chen X. Effective Treatment Therapies of Traditional Chinese Medicine on Hypertension (Chinese version). Guangzhou: Press of Guangzhou; 2003.
37. Li W, Liu L, Puente JG, Li Y, Jiang X, Jin S, et al. Hypertension and health-related quality of life: an epidemiological study in patients attending hospital clinics in China. J Hypertens 2005 Sep;23(9):1667-1676. [doi: 10.1097/01.hjh.0000174971.64589.39] [Medline: 16093911]
38. Pan Y. Treatment and Adjustment with Chinese Medicine and Western Medicine on Hypertension. Hong kong: The Company of Tian Heng Culture Press; 2001.
39. Peng M. Food Therapy of Treatment for Four Seasons on Hypertension (Chinese). Zhengzhou: Press of Peasants in Middle Plain in China; 2004.
40. Liu Z, Yao C. The Nutritious Diet and Food Therapy on Common Chronic Diseases: Hypertension. Beijing: Press of People's Health; 2002.
41. Li J, Xie Y. Nature Therapy of Hypertension (Chinese version). Xian: Press of Shanxi Teaching University; 2005.
42. Hou X. A study on dietary therapy of noodle with high protein to patients with hypertension, hyperlipemia, diabetes (Chinese). Journal of Shangdong Agricultural University 1995;26:445-470 [FREE Full text]
43. Rong W. Treatment using single Semen Cassiae on 43 cases essential hypertension patients (Chinese). Heilongjiang Journal of Traditional Chinese Medicine 2003;4:24-25 [FREE Full text]
44. Tsi D, Das N, Tan B. Effects of aqueous celery (Apium graveolens) extract on lipid parameters of rats fed a high fat diet. Planta Med 1995 Feb 4;61(1):18-21. [doi: 10.1055/s-2006-957990] [Medline: 7700983]
45. Shen C, Pang SMC, Kwong EWY, Cheng Z. The effect of Chinese food therapy on community dwelling Chinese hypertensive patients with Yin-deficiency. J Clin Nurs 2010 Apr;19(7-8):1008-1020. [doi: 10.1111/j.1365-2702.2009.02937.x] [Medline: 20492045]
46. Fiske A, Wetherell JL, Gatz M. Depression in older adults. Annu Rev Clin Psychol 2009 Apr;5(1):363-389 [FREE Full text] [doi: 10.1146/annurev.clinpsy.032408.153621] [Medline: 19327033]
47. Stamler R. Implications of the INTERSALT study. Hypertension 1991 Jan 01;17(1 Suppl):I16-I20. [doi: 10.1161/01.hyp.17.1_suppl.i16] [Medline: 1986996]
48. Cook NEA, Cohen J, Hebert PR, Taylor JO, Hennekens CH. Implications of small reductions in diastolic blood pressure for primary prevention. Arch Intern Med 1995 Apr 10;155(7):701-709. [Medline: 7695458]
49. Zou P, Dennis C, Lee R, Parry M. Dietary Approach to Stop Hypertension with Sodium Reduction for Chinese Canadians (DASHNa-CC): A Pilot Randomized Controlled Trial. J Nutr Health Aging 2017 Dec 9;21(10):1225-1232. [doi: 10.1007/s12603-016-0861-4] [Medline: 29188883]

50. Zou P. Recruitment process of a Chinese immigrant study in Canada. Appl Nurs Res 2017 Aug;36:84-87. [doi: 10.1016/j.apnr.2017.06.005] [Medline: 28720245]
51. Zou P, Dennis C, Lee R, Parry M. Hypertension Prevalence, Health Service Utilization, and Participant Satisfaction: Findings From a Pilot Randomized Controlled Trial in Aged Chinese Canadians. Inquiry 2017 Jan 01;54:46958017724942 [FREE Full text] [doi: 10.1177/0046958017724942] [Medline: 28853303]
52. Glanz K, Rimer B, Viswanath K. Health behavior and health education: Theory, research, and practice. San Francisco, California, United States: Jossey-Bass; 2008.
53. Goyal S, Morita P, Lewis GF, Yu C, Seto E, Cafazzo JA. The Systematic Design of a Behavioural Mobile Health Application for the Self-Management of Type 2 Diabetes. Can J Diabetes 2016 Feb;40(1):95-104. [doi: 10.1016/j.jcjd.2015.06.007] [Medline: 26455762]
54. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. BMJ 2007 Mar 01;334(7591):455-459. [doi: 10.1136/bmj.39108.379965.be]
55. Laws M, Heckscher RA. Racial and ethnic identification practices in public health data systems in New England. Public Health Reports 2002 Jan;117(1):50-61. [doi: 10.1016/s0033-3549(04)50108-5]
56. Mays VM, Ponce NA, Washington DL, Cochran SD. Classification of race and ethnicity: implications for public health. Annu Rev Public Health 2003 Jan;24(1):83-110 [FREE Full text] [doi: 10.1146/annurev.publhealth.24.100901.140927] [Medline: 12668755]
57. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013 Mar 08;15(3):e51 [FREE Full text] [doi: 10.2196/jmir.2350] [Medline: 23475457]
58. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012 Sep;Suppl(s2):49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
59. Macefield R. How To Specify the Participant Group Size for Usability Studies: A Practitioner's Guide. Journal of Usability Studies 2009;5(1):34-45 [FREE Full text]
60. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
61. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8(1):18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
62. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010 Jul 16;10(1):67 [FREE Full text] [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
63. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
64. Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. BMJ 2000;321(7259):504 [FREE Full text] [doi: 10.1136/bmj.321.7259.504] [Medline: 10948038]
65. Wu W, Johnson R, Schepp KG, Berry DL. Electronic Self-report Symptom and Quality of Life for Adolescent Patients With Cancer. Cancer Nursing 2011;34(6):479-486. [doi: 10.1097/ncc.0b013e31820a5bdd]
66. Dwyer J, Picciano MF, Raiten DJ. Collection of food and dietary supplement intake data: What We Eat in America-NHANES. J Nutr 2003 Feb 01;133(2):590S-600S. [doi: 10.1093/jn/133.2.590s]
67. Levitan EB, Wolk A, Mittleman MA. Consistency with the DASH diet and incidence of heart failure. Arch Intern Med 2009 May 11;169(9):851-857 [FREE Full text] [doi: 10.1001/archinternmed.2009.56] [Medline: 19433696]
68. Zhou B, Stamler J, Dennis B, Moag-Stahlberg A, Okuda N, Robertson C, INTERMAP Research Group. Nutrient intakes of middle-aged men and women in China, Japan, United Kingdom, and United States in the late 1990s: the INTERMAP study. J Hum Hypertens 2003 Sep 18;17(9):623-630 [FREE Full text] [doi: 10.1038/sj.jhh.1001605] [Medline: 13679952]
69. Zou P. Traditional Chinese Medicine, Food Therapy, and Hypertension Control: A Narrative Review of Chinese Literature. Am. J. Chin. Med 2016 Dec 06;44(08):1579-1594. [doi: 10.1142/s0192415x16500889]
70. Chahine, Topouchian J, Blacher J, Assemani N, Asmar R, Ibanez I, et al. Validation of four devices: Omron M6 Comfort, Omron HEM-7420, Withings BP-800, and Polygreen KP-7670 for home blood pressure measurement according to the European Society of Hypertension International Protocol. VHRM 2014 Jan:33. [doi: 10.2147/vhrm.s53968]
71. Rao G, Powell-Wiley TM, Ancheta I, Hairston K, Kirley K, Lear SA, et al. Identification of Obesity and Cardiovascular Risk in Ethnically and Racially Diverse Populations. Circulation 2015 Aug 04;132(5):457-472. [doi: 10.1161/cir.0000000000000223]
72. Mason C, Katzmarzyk PT. Effect of the site of measurement of waist circumference on the prevalence of the metabolic syndrome. Am J Cardiol 2009 Jun 15;103(12):1716-1720. [doi: 10.1016/j.amjcard.2009.02.018] [Medline: 19539081]
73. Mason C, Katzmarzyk PT. Variability in waist circumference measurements according to anatomic measurement site. Obesity (Silver Spring) 2009 Sep 02;17(9):1789-1795 [FREE Full text] [doi: 10.1038/oby.2009.87] [Medline: 19343017]
74. Dhaliwal SS, Welborn TA. Measurement error and ethnic comparisons of measures of abdominal obesity. Prev Med 2009 Aug;49(2-3):148-152. [doi: 10.1016/j.ypmed.2009.06.023] [Medline: 19589354]

75. Ware JE. SF-36 health survey update. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3130-3139. [doi: 10.1097/00007632-200012150-00008] [Medline: 11124729]
76. Ware JE, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36). Medical Care 1992;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
77. Dennis CL, Hodnett E, Gallop R, Chalmers B. The effect of peer support on breast-feeding duration among primiparous women: a randomized controlled trial. CMAJ 2002 Jan 08;166(1):21-28 [FREE Full text]
78. Walker R, Morris DW, Greer TL, Trivedi MH. Research staff training in a multisite randomized clinical trial: Methods and recommendations from the Stimulant Reduction Intervention using Dosed Exercise (STRIDE) trial. Addict Res Theory 2014 Dec 18;22(5):407-415 [FREE Full text] [doi: 10.3109/16066359.2013.868446] [Medline: 25379036]
79. Babu GR, Karthik M, Ravi D, Ana Y, Shriyan P, Hasige KK, et al. What makes the pregnant women revisit public hospitals for research? Participant engagement and retention trial in a public hospital (PERTH): an RCT protocol. BMC Pregnancy Childbirth 2018 Sep 12;18(1):369 [FREE Full text] [doi: 10.1186/s12884-018-2000-1] [Medline: 30208868]
80. Popkin BM, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev 2012 Jan;70(1):3-21 [FREE Full text] [doi: 10.1111/j.1753-4887.2011.00456.x] [Medline: 22221213]
81. Perkovic V, Huxley R, Wu Y, Prabhakaran D, MacMahon S. The Burden of Blood Pressure-Related Disease. Hypertension 2007 Dec;50(6):991-997. [doi: 10.1161/hypertensionaha.107.095497]
82. Plaisted CS, Lin P, Ard JD, McClure ML, Svetkey LP. The Effects of Dietary Patterns on Quality of Life. Journal of the American Dietetic Association 1999 Aug;99(8):S84-S89. [doi: 10.1016/s0002-8223(99)00421-6]

Abbreviations

DASH: Dietary Approach to Stop Hypertension
DASHNa-CC: Dietary Approach to Stop Hypertension with Sodium (Na) Reduction for Chinese Canadians
mDASHNa-CC: Mobile Dietary Approach to Stop Hypertension with Sodium (Na) Reduction for Chinese Canadians
SF-36v2: Medical Outcomes Study 36-Item Short-Form version two
TCM: Traditional Chinese Medicine

Edited by H Wu; submitted 25.07.19; peer-reviewed by Y Luo, S Song, H Miller, K Alipasandi; comments to author 01.10.19; revised version received 12.11.19; accepted 09.12.19; published 02.04.20.

Please cite as:
Zou P, Stinson J, Parry M, Dennis CL, Yang Y, Lu Z
A Smartphone App (mDASHNa-CC) to Support Healthy Diet and Hypertension Control for Chinese Canadian Seniors: Protocol for Design, Usability and Feasibility Testing
JMIR Res Protoc 2020;9(4):e15545
URL: https://www.researchprotocols.org/2020/4/e15545
doi:10.2196/15545
PMID:32238343

Protocol

Testing a Real-Time Tenofovir Urine Adherence Assay for Monitoring and Providing Feedback to Preexposure Prophylaxis in Kenya (PUMA): Protocol for a Pilot Randomized Controlled Trial

Paul Drain[1], MD; Kenneth Ngure[2], PhD; Nelly Mugo[3], MBChB; Matthew Spinelli[4], MD; Purba Chatterjee[4], MSc; Peter Bacchetti[4], PhD; David Glidden[4], PhD; Jared Baeten[1], MD, PhD; Monica Gandhi[4], MD, MPH

[1]University of Washington, Seattle, WA, United States

[2]Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

[3]Kenya Medical Research Institute, Nairobi, Kenya

[4]University of California, San Francisco, CA, United States

Corresponding Author:
Monica Gandhi, MD, MPH
University of California
995 Potrero Ave, Bldg 80, W84
San Francisco, CA, 94110
United States
Phone: 1 415 476 4082 ext 127
Email: monica.gandhi@ucsf.edu

Abstract

Background: The worldwide expansion of preexposure prophylaxis (PrEP) with oral tenofovir-disoproxil-fumarate/emtricitabine will be critical to ending the HIV epidemic. However, maintaining daily adherence to PrEP can be difficult, and the accuracy of self-reported adherence is often limited by social desirability bias. Pharmacologic adherence monitoring (measuring drug levels in a biomatrix) has been critical to interpreting PrEP trials, but testing usually requires expensive equipment and skilled personnel. We have recently developed a point-of-care (POC) immunoassay to measure tenofovir in urine, allowing real-time adherence monitoring for the first time.

Objective: The goal of this study is to examine a point-of-care adherence metric in PrEP to support and increase adherence via a randomized controlled trial.

Methods: The paper describes the protocol for a pilot randomized controlled trial to test the acceptability, feasibility, and impact on long-term adherence of implementing a POC urine test to provide real-time adherence feedback among women on PrEP. Eligible women (n=100) will be HIV-negative, ≥18 years old, and recruited from a clinic in Kenya that provides PrEP. Participants will be randomized 1:1 to the intervention of providing real-time feedback via the assay versus standard of care adherence counseling. Acceptability by participants will be assessed by a quantitative survey, as well as by qualitative data collected via in-depth interviews (n=20) and focus group discussions (n=4 groups, 5-10 women each). Feasibility will be assessed by the proportion of women retained in the study, the mean number of missed visits, the proportion of planned urine assessments completed, and messages delivered, while in-depth interviews with providers (n=8) will explore the ease of administering the urine test. Tenofovir levels in hair will serve as long-term adherence metrics. A linear mixed-effects model will estimate the effect of the intervention versus standard of care on logarithmically transformed levels of tenofovir in hair.

Results: This study has been funded by the National Institute of Health, approved by the Kenya Medical Research Institute Institutional Review Board, and will commence in June 2020.

Conclusions: A novel urine assay to measure and deliver information on adherence to PrEP in real-time will be tested for the first time in this trial planned among women on PrEP in Kenya. Study findings will inform a larger-scale trial assessing the impact of real-time adherence monitoring/feedback on HIV prevention. Improving adherence to PrEP will have long-term implications for efforts to end the HIV epidemic worldwide.

Trial Registration: ClinicalTrials.gov NCT03935464; https://clinicaltrials.gov/ct2/show/NCT03935464

International Registered Report Identifier (IRRID): PRR1-10.2196/15029

(JMIR Res Protoc 2020;9(4):e15029) doi:10.2196/15029

KEYWORDS

PrEP; adherence; real-time monitoring and feedback; point of care; trial; Kenya; women; tenofovir; urine test; immunoassay

Introduction

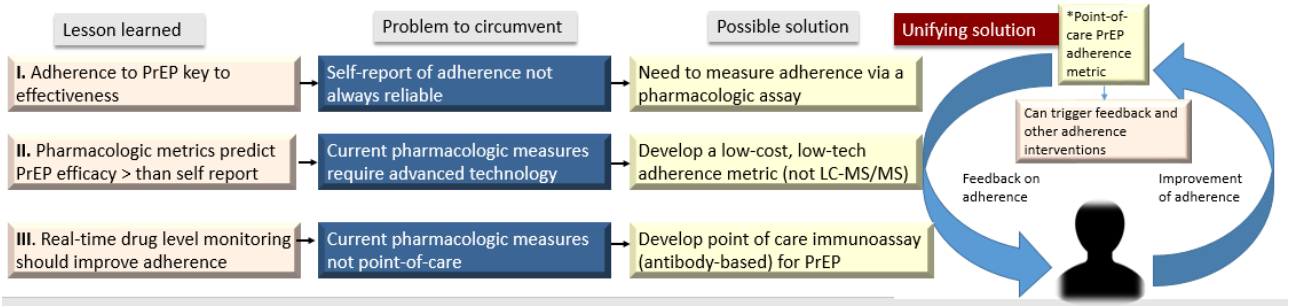
The worldwide expansion of preexposure prophylaxis (PrEP) with oral tenofovir (TFV) disoproxil fumarate/emtricitabine (TDF/FTC) will be critical to ending the HIV epidemic. Oral PrEP has been recommended by the Centers for Disease Control and Prevention [1] and the World Health Organization [2], and is being implemented worldwide. During the early phases of implementation, there have been several challenges and lessons. Importantly, PrEP is only effective for those who are adherent [3], and maintaining daily adherence for prevention can be challenging [4,5].

Daily adherence to PrEP can be difficult to sustain. PrEP was effective in placebo-controlled trials among men-who-have-sex-with-men and transgender women [6], among intravenous drug users [7], and among both men and women in serodiscordant couples [8-10]. However, there was no efficacy of oral PrEP observed in two large trials conducted among young, sexually active women in Africa who were not in serodiscordant relationships [4,5]. In these studies, women in both trials reported >95% adherence to the study drug, but random plasma tenofovir levels among women on the active drug were detectable in fewer than one-third of participants [4,5].

Pharmacologic measures of PrEP drug predict the efficacy of PrEP more accurately than self-reported adherence [4,5,11-15]. In PrEP studies, studies have typically examined the predictive utility of drug concentrations retrospectively using biomatrices such as plasma [16], peripheral blood mononuclear cells [17], hair [18], and dried blood spots [19,20]. Drug levels are usually examined in these biomatrices via liquid chromatography/tandem mass spectrometry (LC-MS/MS). Real-time monitoring of PrEP drug levels, with direct feedback to clients, is difficult to do with LC-MS/MS testing due to the need for specialized equipment and laboratory-based personnel but may improve adherence to oral PrEP [21-24].

Our team has developed a rapid point-of-care (POC) test to objectively assess TFV levels in urine as a measure for PrEP adherence [25]. The immunoassay is highly specific (100%), sensitive (96%), and provided TFV levels in urine that correlated strongly with LC-MS/MS-measured levels (r=0.95) in a study of volunteers administered daily TDF/FTC [25]. In a larger study, where TDF/FTC was administered to HIV-noninfected volunteers at 2, 4, and 7 doses a week, the assay showed the same excellent performance characteristics. From that, we were able to determine an adherence cut-off (in nanograms/milliliter) for TFV in urine for the point-of-care lateral flow assay [26]. We have also shown that low urine TFV levels by this immunoassay predicted future HIV seroconversion events in a large, completed, PrEP demonstration project. [27]. However, this tool has not yet been tested among people on PrEP to determine if real-time monitoring of adherence using a urine assay is feasible, acceptable, and improves PrEP adherence. The principles behind proposing such a trial are summarized in Figure 1.

Figure 1. The rationale behind the PUMA study. PUMA: Point-of-Care Urine Monitoring of Adherence; PrEP: preexposure prophylaxis; LC-MS: liquid chromatography/tandem mass spectrometry.



The objective of this study is to perform a pilot randomized trial in Kenya to test the acceptability, feasibility, and impact on long-term adherence of implementing POC urine TFV testing and providing real-time feedback among women receiving PrEP. We will also conduct a sequential explanatory mixed-methods study to understand both user and provider experiences, preferences, barriers, and facilitators related to POC urine TFV adherence testing after the pilot trial. Our central hypothesis is that real-time PrEP adherence monitoring and feedback, now possible for the first time via a novel point-of-care adherence metric, will motivate adherence and eventually improve the preventative efficacy of PrEP.

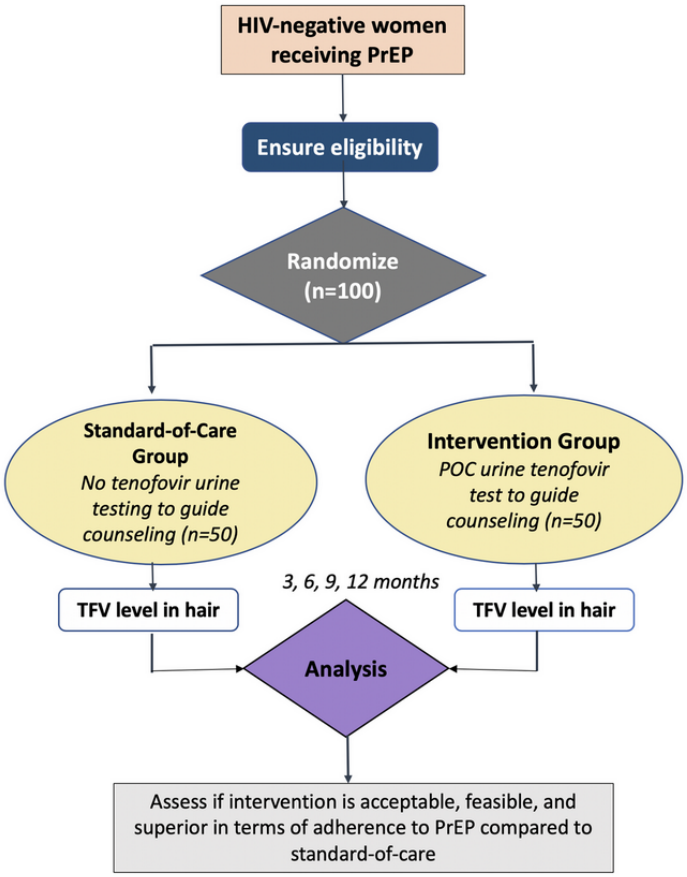
Methods

Study Design

The PUMA (Point-of-Care Urine Monitoring of Adherence) study is an open-label, 12-month, randomized controlled trial in healthy adult women receiving oral PrEP (TDF/FTC). The women receiving PrEP will be randomized (1:1) to the standard

of care (n=50) versus POC urine assay testing (n=50) performed with real-time adherence feedback. Figure 2 describes the study design and the randomization process. Upon randomization, study visits will occur at 3, 6, 9, and 12 months after PrEP initiation.

Figure 2. Schema for the PUMA study. PUMA: Point-of-Care Urine Monitoring of Adherence; PrEP: preexposure prophylaxis; TFV: Tenofovir.



Human Subjects and Informed Consent

The study was filed for ethical approval with the Kenya Medical Research Institute Center for Clinical Research Scientific Committee Meeting (KEMRI/SERU/CCR/0123) and the University of California, San Francisco, Institutional Review Board. All participants will be screened for eligibility and asked to provide written informed consent before study participation. Patients can withdraw from the study at any time or will be withdrawn if HIV antibody testing is positive or if incident HIV infection is detected.

Study Population

A total of 100 women will be recruited, with 50 randomized to the intervention arm and 50 to the standard of care arm (Figure 2). Eligible participants will be women who are HIV-negative, not in a serodiscordant relationship, have an estimated creatinine clearance >60 mL/min, are ≥18 years old, are receiving PrEP, and are returning for a follow-up visit three months after PrEP initiation (which is the second follow-up visit under Kenyan guidelines, with the first occurring at one month after initiation) [28]. We anticipate that randomizing women at their second PrEP follow-up visit, which occurs at three months after PrEP initiation, will help minimize attrition during the study. All people must be able and willing to provide informed consent to participate in the study.

Study Location

The clinical study is being conducted at the Thika Clinic, which is located in an urban center about 40 kilometers outside of Nairobi, Kenya. The clinic is a center of excellence for PrEP delivery in Kenya.

Recruitment, Enrollment, and Randomization

The Thika Clinic has established local recruitment and screening methods that operationalize protocol-specified requirements for eligibility determination in a manner that is tailored to and most efficient for the local setting and target population. Recruitment strategies will include partnering with existing voluntary counseling and testing centers, outreach workers, community organizations (eg, churches), and community mobilization around women's voluntary counseling and testing promotion. Recruitment materials will educate women about PrEP. Screening and enrollment may occur on the same day or may be split across days, depending on the preference of the potential participant.

A member of the study team will approach individuals receiving PrEP. They will describe the study and ask for voluntary participation. Under the 2016 Kenya PrEP guidelines, all people initiating PrEP undergo a clinical assessment to ask about symptoms of acute HIV infection and receive rapid HIV and creatinine testing. Those who want to participate voluntarily

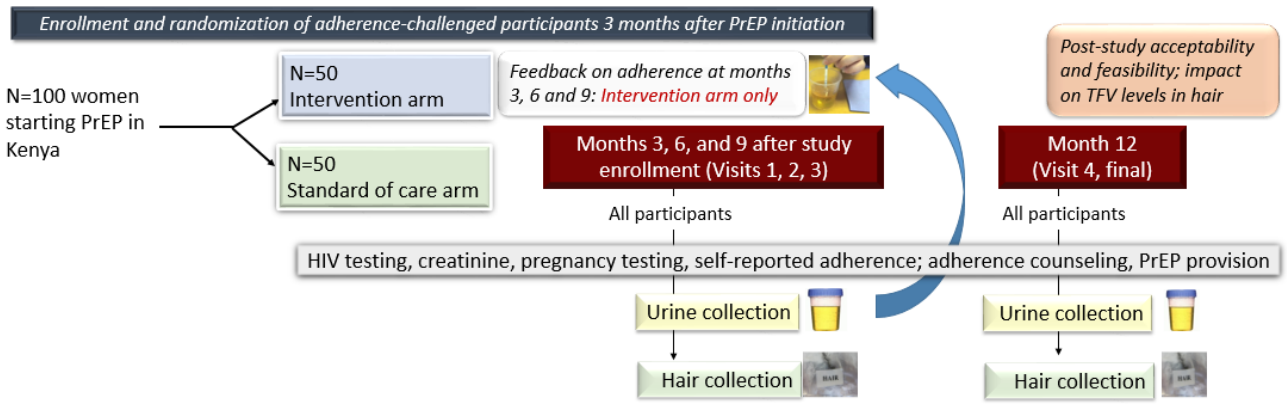
will be taken to a private area of the clinic to be asked several demographic and clinical questions.

After obtaining written informed consent, participants will be randomized (1:1) to receive either POC urine TFV testing with same-day counseling or standard-of-care self-reported adherence monitoring (Figure 2). The study statistician will generate an allocation sequence with random numbers using SAS 9.4 (SAS Institute Inc, Cary, North Carolina, United States). Sequentially numbered, sealed, opaque envelopes containing a study arm allocation and participant identification number will be opened once an eligible, consenting participant is enrolled.

Study Procedures

The study will follow all aspects of Kenya's PrEP guidelines [28] except that those randomized to the intervention arm will receive POC adherence testing by a urine TFV assay (Figure 3). Once written consent is obtained, a member of the study team will collect demographic and health questions related to age, birthdate, income, employment history, prior HIV testing, medical conditions, and current symptoms. They will also obtain each participant's phone number, address, and relevant contact information. After obtaining baseline demographic and clinical data, the study group assignment of the participant will be determined by a random process. A research nurse will then meet the participant in the same clinical exam room and administer a brief clinical questionnaire. The nurse will coordinate the necessary blood draws for the participant so that each participant will only have one blood draw for each clinical visit.

Figure 3. Schedule of evaluations for participants in the intervention and standard of care arms for the PUMA study. PUMA: Point-of-Care Urine Monitoring of Adherence; PrEP: preexposure prophylaxis; TFV: Tenofovir.



People who qualify for and initiate PrEP in Kenya according to the 2016 Kenyan PrEP guidelines [28] are subsequently seen in one month, in three months, and then every three months for repeat HIV testing. Patients on PrEP undergo continued risk assessment and adherence counseling for daily PrEP under the standard of care. PrEP delivery includes measurement of renal function (estimated creatinine clearance >60 mL/min to start PrEP and periodic monitoring over time, aligned to the visit schedule of the study), standard clinical assessment to avoid the continuation of PrEP during acute HIV infection, and adherence counseling. Participants in both arms will be counseled on PrEP and provided with three months of PrEP medication after the one-month visit, according to Kenyan national guidelines.

Follow-Up Visits

At study baseline and months 3, 6, 9, and 12 after enrollment, the study team will collect urine, plasma, whole blood from dried blood spots, and hair samples. Table 1 provides a summary of clinical visits and testing in both study arms. At each study visit, we will offer counseling for participants for HIV testing (pre- and posttesting), HIV infection risk reduction best practices, condom promotion and provision, adherence to PrEP medication, as well as other HIV prevention strategies. Participants in the standard-of-care arm will be directed to the regular clinical waiting area to be seen and evaluated by the study team every three months until study end (Month 12). The study team will prescribe PrEP and additional medications, as well as provide adherence counseling, as appropriate. Kenyan guidelines recommend conducting rapid testing for HIV before dispensing PrEP at each visit. If any participants have a positive rapid test by an oral or blood-based test, confirmatory testing will then be performed, and participants will receive standard HIV care, including initiation of antiretroviral therapy. All participants will receive HIV counseling, condoms, risk reduction counseling, and syndromic management of sexually transmitted infections according to local guidelines.

Table 1. Summary of the clinical visits and laboratory testing for study groups.

Visit requirements	Enrollment	Month 3	Month 6	Month 9	Month 12
Screening and enrollment					
Review eligibility criteria	✓	—	—[a]	—	—
Obtain informed consent	✓	—	—	—	—
Randomization	✓	—	—	—	—
Collect sociodemographic information	✓	—	—	—	—
Research assistant tasks					
Collect and update contact information	✓	✓	✓	✓	✓
Conduct baseline questionnaire	✓	—	—	—	—
Conduct quarterly questionnaire	—	✓	✓	✓	—
Conduct exit study questionnaire	—	—	—	—	✓
Clinical visit by nurse or physician					
Medical history and interval updates	✓	✓	✓	✓	✓
Physical examination (as needed)	✓	✓	✓	✓	✓
Rapid HIV testing	✓	✓	✓	✓	✓
PrEP[a] or drug side effect screen	✓	✓	✓	✓	✓
Collect/assess pill count for prior PrEP	✓	✓	✓	✓	✓
PrEP dispensing for 3-month supply	✓	✓	✓	✓	✓
Adherence and risk reduction counseling	✓	✓	✓	✓	✓
Point-of-care and laboratory testing					
Serum hemoglobin	✓	—	—	—	—
Serum creatinine	✓	—	✓	—	✓
Urinalysis	✓	—	✓	—	✓
POC[b] urine TFV[c] assay (intervention arm)	✓	✓	✓	✓	✓
POC urine TFV assay (standard-of-care arm)	—	—	—	—	✓
Specimen collection and storage					
Stored plasma	✓	✓	✓	✓	✓
Dried blood spot for TFV-DP[d]	✓	✓	✓	✓	✓
Hair sample for TFV-DP	✓	✓	✓	✓	✓

[a]Not applicable.

[b]PrEP: preexposure prophylaxis.

[c]POC: point-of-care.

[d]TFV: tenofovir.

[e]TFV-DP: tenofovir-diphosphate.

Participants in the intervention arm will receive the same treatment as participants in the standard-of-care arm, and also quarterly testing for TFV by a rapid urine diagnostic test. The results of the test will be provided to the participant and will be used to inform enhanced adherence counseling for those who do not demonstrate adequate PrEP adherence. An image of the POC urine adherence assay is shown in Figure 4.

Subsequently, clinic-based POC testing for urine TFV will occur at quarterly visits during the 12-month study period. Counseling messages to the women in the intervention arm will be delivered by the study team, which includes staff with extensive experience in counseling and behavioral interventions. The messages to provide feedback on adherence will be adapted from the HPTN082 (HIV Prevention Trials Network 082) study and other studies that have provided PrEP pharmacologic feedback [29]. The counseling messages for this trial will first be piloted in a small group of women on PrEP at Thika for refinement before actual implementation.

Figure 4. Prototype of point-of-care urine-based tenofovir assay.

Study Outcomes

Participants enrolled in this study will be followed for a total of 12-months from the date of enrollment. The three primary outcomes for this study will be: (1) feasibility of the intervention; (2) acceptability of POC urine tenofovir testing among women receiving PrEP; and (3) preliminary impact on adherence as assessed by a long-term metric of adherence (eg, hair levels).

Feasibility

The feasibility of this intervention will be assessed by the proportion of women retained in the study at 12 months, the mean number of missed visits, the proportion of planned urine assessments completed, and the proportion of messages delivered in the intervention arm. Qualitative data on feasibility will also be assessed via in-depth interviews with research providers at the clinic (n=8) on the acceptability and ease of administering the urine POC test. We will also assess PrEP refills, HIV testing completion, and safety (including accuracy of HIV testing, management of side effects, and social harm). We will also establish via the interviews with providers whether they liked the yes/no assay or would prefer an assay that has more lines indicating high, moderate, or low adherence.

Acceptability

Acceptability will be assessed by a quantitative survey of participants at the end of the study (n=50 in the intervention arm) as well as via qualitative data collected via in-depth interviews with participants (n=20) and focus group discussions (n=4 groups; 5-10 women each). Items to be assessed in the quantitative surveys include questions on feelings about receiving their PrEP adherence results in real-time, likelihood of wanting to receive results of urine testing outside of a study while they are on PrEP, concern about the privacy and security of the data regarding their urine results, grading of the potential impact of knowing their urine TFV results on subsequent medication adherence, and likelihood of taking PrEP just before study visits because they knew the urine test was being conducted. The semistructured interview guide for the qualitative interviews will elicit feelings about the adherence metric and counseling messages, concerns regarding privacy, advantages/disadvantages of receiving such results, and the likely impact of this monitoring test on sustained adherence to PrEP or just short-term adherence. Finally, we will also establish via the interviews whether women liked the yes/no assay or would prefer an assay that has more lines indicating high, moderate, or low adherence.

Long-Term Adherence

Hair levels of TFV and FTC in hair samples will be measured at the 0 (baseline), 3, 6, 9, and 12-month clinic visits after enrollment in the study [18,30,31]. Drug concentrations will be measured at the Hair Analytical Laboratory at the University of California, San Francisco, using validated LC-MS/MS assays [18]. Our methods have been peer-reviewed and approved by the Division of AIDS' Clinical Pharmacology Quality Assurance and Quality Control Program [32], which is based on the US Food and Drug Administration's Guidance for Industry Bioanalytical Method Validation. Hair levels will serve as the efficacy outcome of the pilot trial. Of note, incident HIV infection will be measured but is expected to be low given that individuals will be taking PrEP, and the study would need to be considerably larger to assess incident HIV with much precision. Finally, we will assess genotypic HIV resistance among any seroconverters.

Sample Size Calculation

We assessed sample size based on the primary effectiveness outcome of this pilot randomized trial, an increase in hair levels with the real-time feedback/monitoring in the intervention arm compared to the standard of care arm. For the hair level outcome, we considered the simplified situation of estimating the difference in hair level changes between intervention and control arms using a single postintervention level from each person. In this scenario, and assuming the person-to-person variability in TFV hair levels observed in a similar population (women in Africa not in known, mutually-disclosed serodiscordant couples) in VOICE (Vaginal and Oral Interventions to Control the Epidemic) [5], the 95% confidence interval for an observed difference in means of log (TFV level) would extend 1.5-fold up and down from the estimate with 50 per arm. Given the 7- to 14-fold differences in median hair levels from different adherence levels in various PrEP studies [31], this precision is likely to provide strong enough evidence for improved adherence to justify a subsequent trial of efficacy for preventing HIV infection. Our actual precision will likely be better because of multiple observations per person and the contribution of within-person changes to the overall estimate.

Statistical Analysis

The primary biostatisticians for the study and the Thika Clinic research staff will conduct data management by including procedures to ensure data quality, double data entry, range checks for data values, and encrypting data without patient identifiers for blinded analysis. Participant contact tracing for retention at month 6 and 12 will be performed, particularly to establish PrEP continuation and HIV status. We will use

structured interviews on HIV testing practices and self-reported PrEP adherence (eg, frequency, ability, self-rating, missed doses). We will use REDCap (research electronic data capture) to record all study data.

Our primary analysis of the adherence outcome will be a linear mixed-effects model to estimate the effect of the intervention versus standard-of-care on logarithmically transformed levels of TFV in hair. Since large relative differences between very low hair levels are less clinically important than similar relative differences between higher levels, we will Winsorize [33] undetectable levels to equal the detection limit, and we will add the detection limit to all levels before log transformation. These steps reduce the influence of minor differences between levels at or near the detection limit while preserving the approximate interpretation of back-transformed regression coefficients as fold-effects. The predictor variable for intervention will equal zero for everyone at the baseline (month 0) visit (where hair levels reflect prerandomization baseline) and then will change for those in the intervention arm to equal one at all subsequent visits. A random intercept term will account for within-person correlation across multiple visits. The timepoint will be included as a categorical variable to account for systematic changes over time. We note that inclusion of the preintervention baseline levels will allow the model to account for some of the random person-to-person variation as well as allowing within-person changes to contribute to the estimate of the intervention's effect. Correlations between self-reported adherence over 30 days and hair levels will also be calculated.

Analyses will be by intention-to-treat. Tenofovir levels in hair will be compared between arms using repeated measures mixed-effects models. PrEP discontinuation, defined as missing any refill, will be analyzed as a time-to-event outcome using Cox proportional hazards regression. If the treating clinician discontinues PrEP for safety reasons (but not adherence reasons), follow-up after that will be censored. Adjusted analyses will be done to control for potential confounders based on our prior work assessing correlates of PrEP use: demographics (eg, age, educational level), sexual behaviors (eg, condom use, outside partnerships), medical status (eg, depression), and beliefs (eg, risk perception, PrEP efficacy). SAS Software will be used for analyses.

Results

This study was funded by the National Institutes of Health in September 2018 (R01AI143340) and approved by the Kenya Medical Research Institute Institutional Review Board in August 2019. The study is projected to start in June 2020, with data collection from June 2020 to January 2022, data analysis commencing January 2022, and publication expected by June 2022.

Discussion

A point-of-care test to analyze tenofovir concentrations in urine has recently been developed, and this protocol paper describes the design of the first pilot clinical trial incorporating POC adherence testing and feedback for participants on PrEP. Given the critical contribution of adherence to the preventative efficacy of PrEP, novel interventions to improve long-term PrEP adherence will bolster the ability of this powerful preventative strategy to stem the tide of the HIV pandemic. Self-reported adherence to PrEP is limited by social desirability bias, and no surrogate of PrEP adherence (such as HIV viral loads in those living with HIV) exists. The field has, therefore, increasingly relied on pharmacologic metrics to objectively verify PrEP adherence, but the current methods to measure drug levels are expensive and laboratory-based. The development of a novel antibody-based POC test to monitor PrEP adherence in real-time could be transformative in how we document and intervene against poor PrEP adherence, but this trial will test that concept for the first time.

Prior studies have revealed that real-time monitoring and feedback on adherence could be motivating. In the VOICE study, a placebo-controlled randomized trial testing PrEP in African women, at least 50% of women on the active drug had undetectable TFV in all plasma tested. However, participants over-reported adherence, even at the last study visit where accurate reporting would not affect study participation [5]. Further qualitative work to understand nonadherence in the VOICE study was performed through the Microbicides Trials Network–003D study, which recruited participants with pharmacokinetic (PK) data consistent with low (0% of plasma samples having detectable TFV), inconsistent, or high adherence for retrospective disclosure of plasma TFV results and subsequent capture of reactions [21]. Women with low adherence first expressed surprise at the PK results, then acknowledged they were true and revealed reasons for non-drug-taking during in-depth interviews. Women in all three categories stated that real-time monitoring and feedback would improve adherence and that they would be more honest if presented with objective adherence metrics.

In the iPrEx (Iniciativa Profilaxis Pre-Exposición) open-label extension (OLE) study, a demonstration project of PrEP among men-who-have-sex-with-men and transgender women, drug level testing was performed in plasma and dried blood spots [22]. The results of plasma drug levels over prior weeks were shared with individuals at a later visit, which was reported as highly acceptable [23]. Those with detectable drug in their plasma appreciated receiving validation of adherence, and those without drug detection were not surprised. An in-depth qualitative analysis from iPrEx OLE confirmed the acceptability of drug level feedback and, for those who were not adherent, the motivating effect of receiving such feedback on subsequent adherence. The results of these interviews led authors to conclude that: (1) drug level feedback should be provided as quickly as possible, so new methods should be sought to provide rapid feedback; and (2) drug level feedback will encourage frank adherence discussion and should be provided in the context of adherence counseling.

Urine is a suitable matrix for POC testing for PrEP adherence since urine collection is noninvasive, preferred among youth over blood sampling, and TFV levels in urine correlate with TDF adherence [34-36]. Urine levels reflect plasma levels, which may be particularly useful when daily PrEP adherence is necessary [37]. Plasma and urine PrEP drug levels reflect

only short-term exposure, and thus monitoring via these matrices may be susceptible to "white coat adherence," where adherence improves transiently before a visit. Despite this theoretical concern, this phenomenon has rarely been observed in PrEP delivery [31].

At the study's end, we will have conducted a field test of a novel immunoassay to quantitate TFV in urine as the first POC low-cost adherence metric available for participants on PrEP. The assay will launch a novel modality for adherence intervention. Moreover, the assay could be combined with digital technologies to support PrEP adherence, which are of burgeoning interest in the PrEP adherence intervention field [38,39], especially among youth [40]. Finally, the study findings will inform a larger-scale PrEP trial assessing the impact of real-time adherence monitoring/feedback via the assay. We expect this new TFV POC assay, the metabolite of both TDF and tenofovir alafenamide, to have widespread utility in both HIV prevention and treatment globally. The pilot trial proposed here will examine the acceptability, feasibility, and impact of this novel TFV adherence assay in urine for the first time. Improving adherence to both PrEP and ART will help reduce and prevent HIV transmission in the efforts to end the HIV/AIDS epidemic.

Acknowledgments

We would like to acknowledge the contributions of our collaborators at Alere Rapid Diagnostics (Michael Vincent, Guohong Wang, Warren Rodrigues), who helped develop the assay. This study is supported by a research grant from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, United States (Grant number: R01AI144340, Gandhi PI). The funder of the study (US National Institutes of Health) had no role in the study design, data collection, data analyses, results interpretation, or writing of the report.

Authors' Contributions

MG conceived the concept of the study. PKD and MG drafted the initial protocol manuscript. DG and PB wrote the biostatistical sections. KN, NRM, MAS, PC, JMB participated in the design of the study, reviewed, and edited the protocol manuscript, and approved the final version.

Conflicts of Interest

DVG has accepted fees from Gilead Sciences.

References

1. US Public Health Service. Centers for Disease Control and Prevention. 2018 Apr. Preexposure Prophylaxis for the Prevention of HIV in the United States: A Clinical Practice Guideline - 2017 Update URL: https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2017.pdf [accessed 2018-04-23]
2. World Health Organization. 2018 Dec. Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV: interim guidance URL: http://www.who.int/hiv/pub/guidelines/ARV2018update/en/ [accessed 2019-03-26]
3. Amico KR. Adherence to preexposure chemoprophylaxis: the behavioral bridge from efficacy to effectiveness. Curr Opin HIV AIDS 2012 Nov;7(6):542-548. [doi: 10.1097/COH.0b013e3283582d4a] [Medline: 22964887]
4. Van Damme L, Corneli A, Ahmed K, Agot K, Lombaard J, Kapiga S, et al. Preexposure Prophylaxis for HIV Infection among African Women. N Engl J Med 2012 Aug 02;367(5):411-422. [doi: 10.1056/nejmoa1202614]
5. Marrazzo JM, Ramjee G, Richardson BA, Gomez K, Mgodi N, Nair G, et al. Tenofovir-Based Preexposure Prophylaxis for HIV Infection among African Women. N Engl J Med 2015 Feb 05;372(6):509-518. [doi: 10.1056/nejmoa1402269]
6. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
7. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, Bangkok Tenofovir Study Group. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet 2013 Jun 15;381(9883):2083-2090. [doi: 10.1016/S0140-6736(13)61127-7] [Medline: 23769234]
8. Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, et al. Antiretroviral Preexposure Prophylaxis for Heterosexual HIV Transmission in Botswana. N Engl J Med 2012 Aug 02;367(5):423-434. [doi: 10.1056/nejmoa1110711]
9. Baeten JM, Celum C. Antiretroviral Preexposure Prophylaxis for HIV Prevention. N Engl J Med 2013 Jan 03;368(1):82-84. [doi: 10.1056/nejmc1210464]
10. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral Prophylaxis for HIV Prevention in Heterosexual Men and Women. N Engl J Med 2012 Aug 02;367(5):399-410. [doi: 10.1056/nejmoa1108524]
11. Corneli AL, McKenna K, Perry B, Ahmed K, Agot K, Malamatsho F, et al. The science of being a study participant: FEM-PrEP participants' explanations for overreporting adherence to the study pills and for the whereabouts of unused pills. J Acquir Immune Defic Syndr 2015 Apr 15;68(5):578-584. [doi: 10.1097/QAI.0000000000000525] [Medline: 25761233]

12. van der Straten A, Brown ER, Marrazzo JM, Chirenje MZ, Liu K, Gomez K, MTN-003 VOICE Protocol Team for Microbicide Trials Network. Divergent adherence estimates with pharmacokinetic and behavioural measures in the MTN-003 (VOICE) study. J Int AIDS Soc 2016;19(1):20642 [FREE Full text] [doi: 10.7448/IAS.19.1.20642] [Medline: 26850270]
13. Baker Z, Javanbakht M, Mierzwa S, Pavel C, Lally M, Zimet G, et al. Predictors of Over-Reporting HIV Pre-exposure Prophylaxis (PrEP) Adherence Among Young Men Who Have Sex With Men (YMSM) in Self-Reported Versus Biomarker Data. AIDS Behav 2018 Apr;22(4):1174-1183 [FREE Full text] [doi: 10.1007/s10461-017-1958-4] [Medline: 29079950]
14. Agot K, Taylor D, Corneli AL, Wang M, Ambia J, Kashuba ADM, et al. Accuracy of Self-Report and Pill-Count Measures of Adherence in the FEM-PrEP Clinical Trial: Implications for Future HIV-Prevention Trials. AIDS Behav 2015 May;19(5):743-751 [FREE Full text] [doi: 10.1007/s10461-014-0859-z] [Medline: 25100053]
15. Blumenthal J, Haubrich R. Pre-exposure prophylaxis for HIV infection: how antiretroviral pharmacology helps to monitor and improve adherence. Expert Opin Pharmacother 2013 Sep;14(13):1777-1785 [FREE Full text] [doi: 10.1517/14656566.2013.812072] [Medline: 23800167]
16. Hendrix CW, Andrade A, Bumpus NN, Kashuba AD, Marzinke MA, Moore A, et al. Dose Frequency Ranging Pharmacokinetic Study of Tenofovir-Emtricitabine After Directly Observed Dosing in Healthy Volunteers to Establish Adherence Benchmarks (HPTN 066). AIDS Res Hum Retroviruses 2016 Jan;32(1):32-43 [FREE Full text] [doi: 10.1089/AID.2015.0182] [Medline: 26414912]
17. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, et al. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. Sci Transl Med 2012 Sep 12;4(151):151ra125 [FREE Full text] [doi: 10.1126/scitranslmed.3004006] [Medline: 22972843]
18. Liu AY, Yang Q, Huang Y, Bacchetti P, Anderson PL, Jin C, et al. Strong relationship between oral dose and tenofovir hair levels in a randomized trial: hair as a potential adherence measure for pre-exposure prophylaxis (PrEP). PLoS One 2014;9(1):e83736 [FREE Full text] [doi: 10.1371/journal.pone.0083736] [Medline: 24421901]
19. Castillo-Mancilla JR, Zheng J, Rower JE, Meditz A, Gardner EM, Predhomme J, et al. Tenofovir, emtricitabine, and tenofovir diphosphate in dried blood spots for determining recent and cumulative drug exposure. AIDS Res Hum Retroviruses 2013 Feb;29(2):384-390. [doi: 10.1089/AID.2012.0089] [Medline: 22935078]
20. Anderson PL, Liu AY, Castillo-Mancilla JR, Gardner EM, Seifert SM, McHugh C, et al. Intracellular Tenofovir-Diphosphate and Emtricitabine-Triphosphate in Dried Blood Spots following Directly Observed Therapy. Antimicrob Agents Chemother 2018 Jan;62(1) [FREE Full text] [doi: 10.1128/AAC.01710-17] [Medline: 29038282]
21. van der Straten A, Montgomery ET, Musara P, Etima J, Naidoo S, Laborde N, Microbicide Trials Network-003D Study Team. Disclosure of pharmacokinetic drug results to understand nonadherence. AIDS 2015 Oct 23;29(16):2161-2171 [FREE Full text] [doi: 10.1097/QAD.0000000000000801] [Medline: 26544581]
22. Grant RM, Anderson PL, McMahan V, Liu A, Amico KR, Mehrotra M, iPrEx study team. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. Lancet Infect Dis 2014 Sep;14(9):820-829 [FREE Full text] [doi: 10.1016/S1473-3099(14)70847-3] [Medline: 25065857]
23. Koester KA, Liu A, Eden C, Amico KR, McMahan V, Goicochea P, et al. Acceptability of drug detection monitoring among participants in an open-label pre-exposure prophylaxis study. AIDS Care 2015;27(10):1199-1204 [FREE Full text] [doi: 10.1080/09540121.2015.1039958] [Medline: 26001026]
24. Landovitz RJ, Beymer M, Kofron R, Amico KR, Psaros C, Bushman L, et al. Plasma Tenofovir Levels to Support Adherence to TDF/FTC Preexposure Prophylaxis for HIV Prevention in MSM in Los Angeles, California. J Acquir Immune Defic Syndr 2017 Dec 15;76(5):501-511 [FREE Full text] [doi: 10.1097/QAI.0000000000001538] [Medline: 28902074]
25. Gandhi M, Bacchetti P, Rodrigues WC, Spinelli M, Koss CA, Drain PK, et al. Development and Validation of an Immunoassay for Tenofovir in Urine as a Real-Time Metric of Antiretroviral Adherence. EClinicalMedicine 2018;2-3:22-28 [FREE Full text] [doi: 10.1016/j.eclinm.2018.08.004] [Medline: 30906930]
26. Gandhi M, Bacchetti P, Spinelli MA, Okochi H, Baeten JM, Siriprakaisil O, et al. Brief Report: Validation of a Urine Tenofovir Immunoassay for Adherence Monitoring to PrEP and ART and Establishing the Cutoff for a Point-of-Care Test. J Acquir Immune Defic Syndr 2019 May 01;81(1):72-77. [doi: 10.1097/QAI.0000000000001971] [Medline: 30664078]
27. Spinelli MA, Glidden DV, Rodrigues WC, Wang G, Vincent M, Okochi H, et al. Low tenofovir level in urine by a novel immunoassay is associated with seroconversion in a preexposure prophylaxis demonstration project. AIDS 2019 Apr 01;33(5):867-872. [doi: 10.1097/QAD.0000000000002135] [Medline: 30649051]
28. Guidelines on Use of Antiretroviral Drugs for Treating and Preventing HIV Infections in Kenya. 2016 Mar 01. URL: https://tinyurl.com/s5cswky [accessed 2019-04-20]
29. Celum C, Delany-Moretlwe S, Hosek S, Dye B, Bekker L, Mgodi N, et al. Risk Behavior, Perceptions, and Reasons for PrEP among Young African Women in HPTN 082. 2018 Presented at: 25th Conference on Retroviruses and Opportunistic Infections (CROI); March 4-7; Boston, Massachusetts, United States URL: http://www.croiconference.org/sessions/risk-behavior-perception-and-reasons-prep-among-young-african-women-hptn-082
30. Gandhi M, Murnane PM, Bacchetti P, Elion R, Kolber MA, Cohen SE, et al. Hair levels of preexposure prophylaxis drugs measure adherence and are associated with renal decline among men/transwomen. AIDS 2017 Oct 23;31(16):2245-2251 [FREE Full text] [doi: 10.1097/QAD.0000000000001615] [Medline: 28832411]

31. Koss CA, Bacchetti P, Hillier SL, Livant E, Horng H, Mgodi N, et al. Differences in Cumulative Exposure and Adherence to Tenofovir in the VOICE, iPrEx OLE, and PrEP Demo Studies as Determined via Hair Concentrations. AIDS Res Hum Retroviruses 2017 Aug;33(8):778-783 [FREE Full text] [doi: 10.1089/aid.2016.0202] [Medline: 28253024]
32. DiFrancesco R, Tooley K, Rosenkranz SL, Siminski S, Taylor CR, Pande P, et al. Clinical pharmacology quality assurance for HIV and related infectious diseases research. Clin Pharmacol Ther 2013 Jun;93(6):479-482 [FREE Full text] [doi: 10.1038/clpt.2013.62] [Medline: 23588323]
33. Rivest L. Statistical properties of Winsorized means for skewed distributions. Biometrika 1994;81(2):373-383. [doi: 10.1093/biomet/81.2.373]
34. Marrazzo JM, Scholes D. Acceptability of urine-based screening for Chlamydia trachomatis in asymptomatic young men: a systematic review. Sex Transm Dis 2008 Nov;35(11 Suppl):S28-S33. [doi: 10.1097/OLQ.0b013e31816938ca] [Medline: 18418291]
35. Hadland SE, Levy S. Objective Testing: Urine and Other Drug Tests. Child Adolesc Psychiatr Clin N Am 2016 Jul;25(3):549-565 [FREE Full text] [doi: 10.1016/j.chc.2016.02.005] [Medline: 27338974]
36. Koenig HC, Mounzer K, Daughtridge GW, Sloan CE, Lalley-Chareczko L, Moorthy GS, et al. Urine assay for tenofovir to monitor adherence in real time to tenofovir disoproxil fumarate/emtricitabine as pre-exposure prophylaxis. HIV Med 2017 Jul;18(6):412-418 [FREE Full text] [doi: 10.1111/hiv.12518] [Medline: 28444867]
37. Kearney BP, Ramanathan S, Cheng AK, Ebrahimi R, Shah J. Systemic and renal pharmacokinetics of adefovir and tenofovir upon coadministration. J Clin Pharmacol 2005 Aug;45(8):935-940. [doi: 10.1177/0091270005278949] [Medline: 16027404]
38. Andriesen J, Bull S, Dietrich J, Haberer JE, Van Der Pol B, Voronin Y, et al. Using Digital Technologies in Clinical HIV Research: Real-World Applications and Considerations for Future Work. J Med Internet Res 2017 Jul 31;19(7):e274 [FREE Full text] [doi: 10.2196/jmir.7513] [Medline: 28760729]
39. Sullivan PS, Driggers R, Stekler JD, Siegler A, Goldenberg T, McDougal SJ, et al. Usability and Acceptability of a Mobile Comprehensive HIV Prevention App for Men Who Have Sex With Men: A Pilot Study. JMIR Mhealth Uhealth 2017 Mar 09;5(3):e26 [FREE Full text] [doi: 10.2196/mhealth.7199] [Medline: 28279949]
40. Biello KB, Marrow E, Mimiaga MJ, Sullivan P, Hightow-Weidman L, Mayer KH. A Mobile-Based App (MyChoices) to Increase Uptake of HIV Testing and Pre-Exposure Prophylaxis by Young Men Who Have Sex With Men: Protocol for a Pilot Randomized Controlled Trial. JMIR Res Protoc 2019 Jan 07;8(1):e10694 [FREE Full text] [doi: 10.2196/10694] [Medline: 30617042]

Abbreviations

HPTN082: HIV Prevention Trials Network 082
iPrEx: Iniciativa Profilaxis Pre-Exposición
LC-MS/MS: liquid chromatography/tandem mass spectrometry
MTN: Microbicides Trials Network
OLE: open-label extension
POC: point of care
PrEP: preexposure prophylaxis
PK: pharmacokinetic
PUMA: Point-of-Care Urine Monitoring of Adherence
REDCap: Research Electronic Data Capture
TDF/FTC: tenofovir disoproxil fumarate/emtricitabine
TFV: tenofovir
VOICE: Vaginal and Oral Interventions to Control the Epidemic

Edited by G Eysenbach; submitted 13.06.19; peer-reviewed by T Torres, L Wang, R Giguere; comments to author 05.08.19; revised version received 23.08.19; accepted 28.08.19; published 02.04.20.

Please cite as:
Drain P, Ngure K, Mugo N, Spinelli M, Chatterjee P, Bacchetti P, Glidden D, Baeten J, Gandhi M
Testing a Real-Time Tenofovir Urine Adherence Assay for Monitoring and Providing Feedback to Preexposure Prophylaxis in Kenya (PUMA): Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2020;9(4):e15029
URL: https://www.researchprotocols.org/2020/4/e15029
doi:10.2196/15029
PMID:32238341

Protocol

Improving the Use of Simulation in Nursing Education: Protocol for a Realist Review

Torbjørg Træland Meum[1], PhD; Åshild Slettebø[1], PhD; Mariann Fossum[1], PhD

Department of Health and Nursing Science, University of Agder, Grimstad, Norway

Corresponding Author:
Torbjørg Træland Meum, PhD
Department of Health and Nursing Science
University of Agder
Box 509
Grimstad, 4898
Norway
Phone: 47 94888747
Email: torbjorg.t.meum@uia.no

Abstract

Background: Nursing education has evolved in line with societal needs, and simulation-based learning (SBL) is increasingly being used to bridge the gap between practice and education. Previous literature reviews have demonstrated the effectiveness of using SBL in nursing education. However, there is a need to explore how and why it *works* to expand the theoretical foundation of SBL. Realist reviews are a theory-based approach to synthesizing existing evidence on how complex programs work in particular contexts or settings.

Objective: This review aims to understand how, why, and in what circumstances the use of simulation affects learning as part of the bachelor's program in nursing.

Methods: A realist review will be conducted in accordance with the realist template for a systematic review. In particular, we will identify and explore the underlying assumption of how SBL is supposed to work, that is, identify and explore program theories of SBL. The review will be carried out as an iterative process of searching, appraising, and synthesizing the evidence to uncover theoretical concepts that explain the causal effects of SBL. In the final section of the review, we will involve stakeholders in the Norwegian community in a web-based Delphi survey to ensure that the emerging theoretical framework derived from the published literature aligns with stakeholders' experience in practice.

Results: The Norwegian Centre for Research Data (project number 60415) has approved the study. We have performed an initial literature search, whereas quality appraisal and data extraction are ongoing processes.

Conclusions: The final outcome of the review is anticipated to extend the theoretical foundation for using simulation as an integrated component of the bachelor's program in nursing. Furthermore, the findings will be used to produce a briefing document containing guidance for national stakeholders in the community of simulation-based nursing education. Finally, the review findings will be disseminated in a peer-reviewed journal as well as national and international conferences.

International Registered Report Identifier (IRRID): DERR1-10.2196/16363

(JMIR Res Protoc 2020;9(4):e16363) doi:10.2196/16363

KEYWORDS

education; nursing; learning; realist review; simulation training

Introduction

The health care sector has undergone radical changes over the last few decades, a fact that has also affected nursing education. Demographic changes, technological development, and innovation in health science have led to increasing requirements being placed on the education of health professionals. A study carried out by the Carnegie Foundation for the Advancement of Teaching [1] has explored the strengths and weaknesses of nursing education, and the authors call for a radical transformation of nursing education. In particular, they describe a gap between education and practice, that is, "the ability of a practice setting to adopt and reflect what was taught in academic institutions" [1]. The authors also highlight the practice-education gap "as it is becoming more and more

difficult for nursing education to keep pace with the rapid changes in clinical practice driven by research and new technologies" [1]. Patient safety is the cornerstone of high-quality health care [2]. The Institute of Medicine has highlighted evidence-based practice as one of the key competences all clinicians should possess to meet the needs of 21st-century health care services [3]. These competences have been adapted to the nursing community, which has in turn proposed statements on the knowledge, skills, and attitudes that should be developed during prelicensure nursing education [4]. The ability to integrate scientific knowledge into practice requires good clinical judgment and reasoning, which is considered an essential part of nursing. Clinical judgment in nursing has been defined as "interpretation or conclusion about a patient's needs, concerns, or health problems, and/or the judgement to take action (or not), use or modify standard approaches, or improvise new ones as deemed appropriate by the patient's response" [5]. From this point of view, clinical reasoning refers to the processes by which nurses and other clinicians arrive at their judgment, including analytical processes, intuition, and narrative thinking [5]. Similar research has also pointed to the situated, practice-based aspect of clinical reasoning and emphasized the ability to discern the relevance of the evidence behind general scientific knowledge and how it applies to a particular clinical situation [6]. The rapid expansion of clinical knowledge over the past decades has led to increasing demands on health professionals to employ evidence-based practice [3]. Clinical judgment is a capability that evolves in line with clinical experience, and the challenge is to integrate clinical knowledge into classroom teaching [7].

Simulation-based learning (SBL) has in recent years been increasingly utilized to bridge the gap between practice and education. The International Nursing Association for Clinical Simulation and Learning has defined simulation as "an educational strategy in which a particular set of conditions are created or replicated to resemble authentic situations that are possible in real life" [8]. The use of simulation in the education of health professionals is not a new innovation; however, the evolving use of advanced technology in recent decades has provided new capabilities for the educational use of simulation. Some of the early versions of full-body patient simulators were designed for subject-specific areas. For example, a Norwegian company developed Resuci Anne in 1960 for cardiopulmonary resuscitation training [9,10]. Since then, full-scale patient simulators that facilitate the formation of dynamic patient situations that fully mirror actual clinical settings have become available [11]. Accordingly, a number of simulation modalities, such as full- and part-body models with low- and high-tech features, computer-based programs, and standardized patients, have been used as an educational resource in health care [11]. Since then, simulation has gradually been integrated into the education of health professionals as an educational intervention that includes digital technologies, human resources, and educational strategies. Experiential learning [12] and situated learning theory [13] have inspired the design and development of simulated learning activities in nursing [14]. These theories emphasize active engagement in the learning process, in which knowledge is created through the transformation of experience. All in all, advancement in educational principles and digital technology has provided the opportunity to create realistic learning activities in a safe environment and, thus, bridge the gap between the classroom and clinical practice.

A growing body of evidence in the simulation literature has demonstrated the effectiveness of simulation in the teaching of clinical knowledge, procedural skills, and teamwork [11,15]. A systematic literature review by Lapkin et al [16] showed that the use of patient simulation manikins in teaching improved knowledge acquisition, critical thinking, and the ability to identify deteriorating patients. Similarly, a recent literature review conducted by Cant and Copper [17] revealed positive outcomes of simulation education for knowledge acquisition, psychomotor skills, self-efficiency, satisfaction, confidence, and critical thinking. Virtual patient simulation is also increasingly used in nursing education. A randomized controlled trial by Liaw et al [18] compared virtual simulation with manikin-based simulation and demonstrated that both the simulation modalities were effective learning strategies for improving nursing students' clinical performance. Although previous research has shown promising effects for using simulation as an educational intervention, studies have also called for a deeper theoretical understanding of how and why simulation contributes to learning [19-22].

The realist review is a method that is increasingly used in health care to explore "What works for whom, in what circumstances and why?" [23-25]. The method is based on a realist philosophy of science, which positions itself between positivism and constructivism [23]. Basically, a realist approach recognizes the existence of an external social reality. However, there is a social reality that cannot be measured directly because our knowledge of it is processed through our brains, human senses, language, and culture. The realist review was developed as a theory-based approach to synthesizing existing evidence on how complex programs work in particular contexts or settings. The unit of analyses is to identify, test, and refine the program theory or theories, that is, the underlying assumptions about how an intervention is supposed to work [24]. The analytical process in realist review emphasizes the causal relationship between contexts, mechanisms, and outcomes (CMO). This process is described as the CMO configuration where the concept of mechanism is considered as the causal force that operates in a particular context to generate the desired outcome. All in all, the analytical process involves iterative testing of the CMO configurations and refinement of the theoretical concepts derived from relevant evidence (qualitative, quantitative, and mixed method studies) [24].

Realist reviews have been used in health care as an approach to the synthesis of evidence in various disciplines such as implementation science [26] and internet-based medical education [27]. Similarly, McGaghie et al [28] used a combined critical and realist review methodology to evaluate research on simulation-based education and highlighted 12 features and best practices to promote the educational impact of using simulation. Research on SBL in nursing education is extensive; however, realist reviews have rarely been used in this research community. Accordingly, we consider a realist review to be a promising method to uncover the causal effects of SBL and address the

call for more theory-driven research to promote the educational use of simulation in nursing education.

Methods

The Purpose of the Review

The overall purpose of this review is to gain a deeper theoretical insight into how the use of simulation affects learning to promote the educational use of simulation as an integral part of the bachelor's program in nursing. It is part of a larger research project focusing on professional development at the department of nursing of a large regional university in Norway. Moreover, the project is affiliated with a research group at the department of nursing science (health care services, ethics, and quality) that highlights innovative and practice-based research to educate future-oriented nurses. The cocreation of knowledge is one of this university's overarching visions as it strives to develop and implement future-oriented, varied, student-centered, and practice-based teaching and learning methods at all levels of instruction. In accordance with this vision, SBL has for many years been an established method of learning as a supplement to traditional teaching. On the basis of challenges, such as an increasing intake of students, it is anticipated that SBL will be used to an even greater extent as an integrated part of the curriculum.

Research Aim

This review aims to identify how the use of simulation affects learning in the bachelor's program in nursing.

Research Question

The research conducted in this study will explore what sort of SBL "works" for whom and in what circumstances. The study will supplement and expand upon previous research in this field, exploring the following questions:

1. How does the use of SBL affect the development of clinical reasoning, clinical judgment, and skills in nursing education?
2. How are the principles of learning translated into learning activities in the simulation setting?
3. What characterizes conditions in the simulation environment that facilitate learning in the bachelor's program in nursing?

Objectives

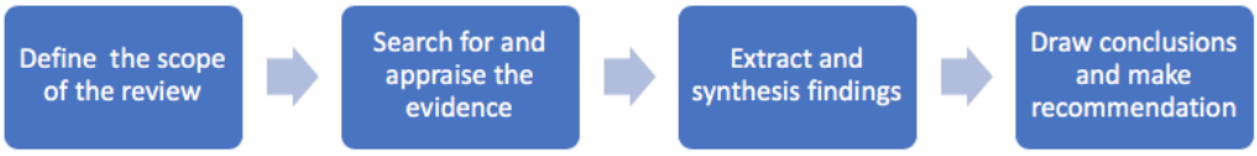
The research objectives of the study are as follows:

1. To conduct a realist review to understand how, why, and in what circumstances the use of simulation affects learning as part of the bachelor's program in nursing.
2. To identify and refine program theories for SBL.
3. To make recommendations with regard to how the potential for using simulation may be utilized in the bachelor's program in nursing.

Study Design

A systematic literature review inspired by the realist approach to evaluation will be carried out [29]. The use of a realist review is a relatively new strategy for synthesizing research and methodological guidance, and training materials have been developed by an interdisciplinary research team [24]. The review process mainly follows the same steps as traditional systematic reviews. However, the unit of analysis in a realist approach is the refinement of program theory, which is characterized as an iterative process that includes both qualitative and quantitative studies [25]. Accordingly, we will utilize the realist template for systematic review [25] and follow the steps illustrated in Figure 1.

Figure 1. Review design.



Stage 1: Defining the Scope of the Review

The first step is to refine the purpose of the review and identify candidate theories, that is, the theoretical basis for why SBL works [24,27]. Two researchers will take part in every step of the process, and one of the team members has extensive knowledge of SBL. On the basis of our knowledge of previous research on the use of SBL, we have acknowledged the need to explore how, why, and for whom it works. Our initial research questions will be followed with a special emphasis on *what sorts of SBL work for whom and in what circumstances.* The objective is to elucidate what and how SBL promotes clinical judgment and clinical skills in the bachelor's program in nursing and then explain this elucidation through the use of middle-range theory. A key point at this stage is to identify existing theories regarding the use of SBL in the bachelor's program in nursing. Previous research has identified Jeffries Framework [30], Kolb's theory of experiential learning [12], and Bandura's social cognitive theory [31] as the most frequently used frameworks and theories in SBL [22]. These theories are characterized as middle-range theories (theories that can usefully be applied to a family of interventions) [24]. These theories, therefore, provide a starting point in the process of identifying program theories that strive to explain the chain of events underlying SBL. As this is an iterative process, during the search and screening stages, we will attempt to expand the list of relevant middle-range theories.

Stage 2: Searching For and Appraising the Evidence

The search strategy will follow the guidelines provided by Booth et al [32] and include six relevant web-based databases: Cochrane Library, Cumulative Index of Nursing and Allied

Health Literature, Medical Literature Analysis and Retrieval System Online, EMBASE, Education Resources Information Center, and Web of Science. The search for evidence will be carried out in collaboration with a health science librarian. Furthermore, it will mainly include the following search terms: "Nursing students," Nursing education," Baccalaureate," "Simulation," "Simulated environment," "Simulation training," "Manikin," and "Anatomic model." Each search will use the relevant search terms or MESH/thesaurus/keyword heading for each database. As indicated by our research questions, we will focus on outcomes related to clinical judgment, clinical reasoning, and clinical skills. However, we have not included these terms in our database search because of their ambiguity, which may lead to the exclusion of relevant studies. To ensure sensitivity, we will do a manual screening with these terms as part of the inclusion criteria. SBL is a comprehensive field of study, and we need to make some pragmatic decisions for which studies we want to include and exclude. Simulation in health care is usually classified in terms of fidelity, that is, the level of realism associated with a particular simulation activity [33]. However, the concept of "fidelity" is not clearly stated in the literature [34], and we will use the definitions drawn up by the Society for Simulation in Healthcare to distinguish between the physical, psychological, and environmental aspects of fidelity [33]. Recent literature reviews have mainly focused on high-fidelity simulation in laboratory settings using computerized full-body manikins [35]. Although high-fidelity simulators are considered a useful learning resource, McGaghie et al [28] have emphasized the link between educational goals and simulation tools as a key principle for the effective use of simulation. Thus, we will include medium- and low-fidelity simulators in the initial screening as we consider these to be the most appropriate in teaching basic nursing. In addition, we want to identify learning outcomes using Kirkpatrick's 4 levels of evaluation: (1) reaction, (2) learning, (3) behavior, and (4) results [36]. Accordingly, our initial inclusion and exclusion criteria are shown in Textboxes 1 and 2.

Textbox 1. Inclusion criteria for this study.

- Bachelor's degree in nursing
- Focus on the student's perspective
- Medium- and low-fidelity simulators
- Every phase of the simulation process included
- Clinical judgment, clinical reasoning, and clinical skills
- Learning outcome related to at least level 2 of the training evaluation described by Kirkpatrick
- Peer-reviewed research paper (qualitative, quantitative, and mixed method studies) written in English
- Published between 2014 and 2019

Textbox 2. Exclusion criteria for this study.

- Continuing education (nurse practitioner, advanced nursing, and midwife)
- High-fidelity simulators, serious games, electronic learning, and virtual reality
- Comparison of different simulation methods
- Interdisciplinary simulation
- Disaster management
- Review articles and doctoral thesis

The preliminary criteria have been developed for searching and identifying middle-range theories that explain *how* and *why* SBL works (or does not work). In accordance with the recommendations for realist reviews, we will include qualitative, quantitative, and mixed method studies. However, we will not conduct snowball sampling or include "grey" literature such as reports, theses, or conference papers.

Search results will be saved as text files and downloaded into a web-based software platform (Covidence) for screening and quality appraisal. Two members of the review team will screen the title and abstract with respect to the inclusion and exclusion criteria. A full-text screening and quality assessment will then be performed based on quality appraisal criteria for realist reviews, that is, relevance (whether it can contribute to theory building and/or testing) and rigor (whether the methods used to generate the relevant data are credible and trustworthy) [24]. In addition, we will use the criteria described by Dixon-Woods et al [37] as well as the Mixed Methods Appraisal Tool [38] to aid our decision making on methodological credibility. When agreement is reached, data will be converted to a flowchart to illustrate the search and screening process as well as the final selection of included papers. References and data derived from the quality assessment will first be stored in Covidence and then transferred to a Microsoft Excel spreadsheet.

Stage 3: Data Extraction and Synthesizing the Results

At this stage, we will transfer both papers and bibliographical data to NVivo software (QSR International) for further data extraction and syntheses. First and foremost, the characteristics

of the included studies will be listed in a table. One of the advantages of NVivo is its ability to transfer bibliographical data to a classification sheet when a library is imported from a reference management software program, for example, Mendeley. Descriptive information on the included studies (titles, authors, sources, and publication year) will, thus, be transferred to the classification sheet in NVivo, which will form the starting point for further data extraction. On the basis of previous realist reviews [39], we will add data domains to the classification sheet related to research design, educational setting, educational consequences, and outcome measures. Furthermore, we will use NVivo to code section of texts that may prove useful for constructing theory and refining theoretical concepts. Generative mechanisms (mechanism of action) are considered to be the key unit of analysis, and data coding will involve identifying the interrelations between CMO in the included studies. Data extraction and synthesis are considered to be an interwoven process where "raw data" captured in the included studies will be used to make sense of the causal relationship between CMO identified in the included studies.

Data synthesis will follow the realist review guidelines through using an interpretive approach to data synthesis [24]. To make sense of our "raw" data, we will incorporate the concepts identified in the primary studies into a higher-order theoretical structure [40]. First of all, we will uncover any semirecurring patterns of behavior (demi-regularities) that may be present in the included studies. We will then explore if our initial middle-range theories are able to explain why these demi-regularities emerge under the contexts reported in our included papers [39]. The middle-range theories will be treated as theoretical data and included in a constant comparison of key demi-regularities derived from the included studies with emerging theoretical conceptualization [24]. The main objective of the review process is to refine the program theory, which may involve several iterations that include testing and refining to progress toward a refined theoretical framework of SBL in nursing education.

Stage 4: Drawing Conclusions and Making Recommendations

The findings will provide theoretical and practical implications for SBL in nursing education. The inductive analysis and synthesis of data extracted from the included studies have the potential to generate theoretical concepts that explain the causal effect of SBL. Uncovering the generative mechanisms provides us with the ability to form a preliminary theoretical framework in our efforts to refine the program theory of SBL [29]. In line with similar reviews, the findings will be transformed into recommendations that can be used to inform policy makers and practitioners. The intended outcome is that practitioners will take note of the findings and implement them. Thus, we expect the review's findings to influence the design of new programs, at this stage involving practitioners. Stakeholders' practical knowledge will be used in the refinement of the program theory to ensure that the emerging theoretical framework derived from the published literature aligns with stakeholders' experiences in practice.

Stakeholders will mainly be involved through using a web-based Delphi method. This is a well-known method for giving structure to group processes used to identify problems, setting goals and priorities, and identifying problems as well as solutions [41]. The method has evolved into a group-facilitating technique in health and social care. It contains an iterative multistage process designed to transform opinion into group consensus [42]. The method's democratic, structured approach that utilizes participants' collective knowledge is considered to be a key advantage [43], having the potential to broaden the knowledge within the nursing profession [42]. The Delphi method has been used in various settings for different purposes; in this particular case, we will apply the research-based guidelines outlined by Hasson et al [42] and Okoli and Pawlowski [44] to maintain validity and credibility in the research process. Selecting qualified experts (a panel of informed individuals) plays a key role in this process. Therefore, our preparations will include using an Excel spreadsheet as a tool to categorize and identify relevant experts, that is, knowledgeable and experienced professionals in SBL. This study will include the Norwegian community, and we will attempt to get input from a wide range of simulation experts comprising associate professors, researchers, and members of professional networks. Furthermore, we will identify relevant panelists from various organizational units such as universities, the Norwegian Nurses Organisation, and "MedSimNorge" (the national network for simulation in the health care sector). Potential participants will be identified by gatekeepers in relevant organizations, and we aim to recruit about 35 participants to the panel. Furthermore, the literature synthesis' key findings will be transformed into statements intended as briefing material for the panelists. In accordance with the guidelines for the Delphi technique, we will carry out several rounds of online questionnaires (brainstorming, narrowing, and ranking) until consensus is reached [44]. The briefing document, as well as three or four additional questions, will be sent to the participants in the process' first round. Subsequent rounds will involve validating, refining, and ranking the statements from previous rounds. The goal of the final phase is to reach consensus about ranking relevant statements, and the panelist will be asked to rank each potential item on a 7-point Likert scale (1=strongly disagree to 7=strongly agree). The ranking results will then be collated and inserted into an Excel spreadsheet and analyzed using nonparametric statistics (Kendall W). In addition, any free-text comments from the panelist will be analyzed thematically. The findings from the Delphi survey will be summarized, a process that involves incorporating the stakeholders' points of view into the final part of the analytical process.

The steps in the review process are overlapping and iterative, which means that the review steps can be revised throughout the process as new ideas and evidence emerge. We will, thus, follow an iterative process until theoretical saturation is achieved, that is, when new data do not provide any new theoretical insight into the emerging theory. Finally, the study will be reported in accordance with publication standards for realist review [45]. A diagram will be used to present the final program theory.

Ethics and Dissemination

This study has been approved by the Norwegian Centre for Research Data (project number 60415). The review findings will be presented at national and international conferences, including the annual conference organized by the Society for Simulation in Europe. Finally, the findings will be disseminated in a peer-reviewed journal.

Results

The initial search generated 4830 unique references, and after initial screening, we have included 113 studies. Data extraction and synthesis are ongoing processes, and we plan to complete a first draft of the literature review in June 2020.

Discussion

Simulation is still a relatively new field of research, and we expect that the findings from this realist review will lead to theoretical and practical implications for the use of simulation in nursing education. Several studies have emphasized the need for conceptual and theoretical frameworks for the use of simulation [14,19], and the final outcome of this review is expected to generate theoretical concepts that may explain the causal effects of simulation as an integrated part of nursing education. The findings will then be used to produce a briefing document to guide practitioners in designing educational programs for the bachelor's degree in nursing. Although the intended audience for this review is nursing educators, we expect the findings to be relevant to other health care professionals.

In this study, we have outlined the different steps involved in a realist review, and we consider it as a promising method to unpack the complexity of SBL. The realist review method has increasingly been used across different research fields; however, it has rarely been used in simulation-based nursing education. A major advantage of the method is to move beyond the effect of using simulation toward exploration of the causal mechanisms involved in the intervention. However, there are some key issues to consider when applying the method. As previously mentioned, the realist review method is based on a realist philosophy of science [29]. Thus, it is necessary to consider the underlying ontological and epistemological assumptions of the realist philosophy to understand the methodological implications. The realist approach is regarded as a middle way between positivism and interpretivism and embraces a variety of qualitative and quantitative methods [46]. This makes it particularly appropriate for the study of SBL that incorporates aspects of both natural science and social science. A central theme of a realist approach is the power of generative mechanisms; however, it can be challenging to know how to identify and define them. To gain a broader understanding of these issues, we will incorporate recent research from other research disciplines that deal with methodological implications of critical realism [46]. We, thus, expect the study to increase the insight into methodological challenges that have been pointed out in similar research [39].

Acknowledgments

This publication is supported by the Department of Health and Nursing Science, University of Agder.

Authors' Contributions

TM designed and drafted the protocol with input from ÅS and MF. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Benner P, Sutphen M, Leonard V, Day L. Educating Nurses: A Call For Radical Transformation. San Francisco: Jossey-bass; 2010.
2. Hughes RG. Patient Safety and Quality: An Evidence-based Handbook for Nurses. Rockville: Agency for Healthcare Research and Quality; 2008.
3. Greiner A, Knebel E. Health Professions Education: A Bridge to Quality. Washington, DC: National Academies Press; 2003.
4. Cronenwett L, Sherwood G, Barnsteiner J, Disch J, Johnson J, Mitchell P, et al. Quality and safety education for nurses. Nurs Outlook 2007;55(3):122-131. [doi: 10.1016/j.outlook.2007.02.006] [Medline: 17524799]
5. Tanner CA. Thinking like a nurse: a research-based model of clinical judgment in nursing. J Nurs Educ 2006 Jun;45(6):204-211. [doi: 10.3928/01484834-20060601-04] [Medline: 16780008]
6. Benner P, Hughes RG, Sutphen M. Clinical reasoning, decisionmaking, and action: thinking critically and clinically. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: Agency for Healthcare Research and Quality (US); 2008:87-110.
7. Benner P. Curricular and pedagogical implications for the Carnegie Study, educating nurses: a call for radical transformation. Asian Nurs Res (Korean Soc Nurs Sci) 2015 Mar;9(1):1-6 [FREE Full text] [doi: 10.1016/j.anr.2015.02.001] [Medline: 25829203]
8. INACL Standards Committee. INACSL standards of best practice: SimulationSM Simulation glossary. Clin Simul Nurs 2016;12:S39-S47. [doi: 10.1016/j.ecns.2016.09.012]

9. Buck GH. Development of simulators in medical education. Gesnerus 1991;48 Pt 1:7-28. [Medline: 1855669]
10. Hayden JK, Smiley RA, Alexander M, Kardong-Edgren S, Jeffries PR. The NCSBN National Simulation Study: a longitudinal, randomized, controlled study replacing clinical hours with simulation in prelicensure nursing education. J Nurs Regul 2014;5(2):S3-40. [doi: 10.1016/s2155-8256(15)30062-4]
11. Nagle BM, McHale JM, Alexander GA, French BM. Incorporating scenario-based simulation into a hospital nursing education program. J Contin Educ Nurs 2009 Jan;40(1):18-25; quiz 26. [doi: 10.3928/00220124-20090101-02] [Medline: 19226995]
12. Kolb AY, Kolb DA. Learning styles and learning spaces: enhancing experiential learning in higher education. Acad Manag Learn Educ 2005;4(2):193-212. [doi: 10.5465/amle.2005.17268566]
13. Lave J, Wenger E. Situated Learning: Legitimate Peripheral Participation. Cambridge: Cambridge University Press; 1991.
14. Bland AJ, Topping A, Wood B. A concept analysis of simulation as a learning strategy in the education of undergraduate nursing students. Nurse Educ Today 2011 Oct;31(7):664-670. [doi: 10.1016/j.nedt.2010.10.013] [Medline: 21056920]
15. Okuda Y, Bryson EO, DeMaria S, Jacobson L, Quinones J, Shen B, et al. The utility of simulation in medical education: what is the evidence? Mt Sinai J Med 2009 Aug;76(4):330-343. [doi: 10.1002/msj.20127] [Medline: 19642147]
16. Lapkin S, Levett-Jones T, Bellchambers H, Fernandez R. Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduate nursing students: a systematic review. Clinical Simulation in Nursing 2010 Nov;6(6):e207-e222. [doi: 10.1016/j.ecns.2010.05.005]
17. Cant RP, Cooper SJ. Use of simulation-based learning in undergraduate nurse education: an umbrella systematic review. Nurse Educ Today 2017 Feb;49:63-71. [doi: 10.1016/j.nedt.2016.11.015] [Medline: 27902949]
18. Liaw SY, Chan SW, Chen FG, Hooi SC, Siau C. Comparison of virtual patient simulation with mannequin-based simulation for improving clinical performances in assessing and managing clinical deterioration: randomized controlled trial. J Med Internet Res 2014 Sep 17;16(9):e214 [FREE Full text] [doi: 10.2196/jmir.3322] [Medline: 25230684]
19. Issenberg SB, Ringsted C, Ostergaard D, Dieckmann P. Setting a research agenda for simulation-based healthcare education: a synthesis of the outcome from an Utstein style meeting. Simul Healthc 2011 Jun;6(3):155-167. [doi: 10.1097/SIH.0b013e3182207c24] [Medline: 21642804]
20. Graham AC, McAleer S. An overview of realist evaluation for simulation-based education. Adv Simul (Lond) 2018;3:13 [FREE Full text] [doi: 10.1186/s41077-018-0073-6] [Medline: 30026966]
21. Bajpai S, Semwal M, Bajpai R, Car J, Ho AH. Health professions' digital education: review of learning theories in randomized controlled trials by the digital health education collaboration. J Med Internet Res 2019 Mar 12;21(3):e12912 [FREE Full text] [doi: 10.2196/12912] [Medline: 30860483]
22. Lavoie P, Michaud C, Bélisle M, Boyer L, Gosselin É, Grondin M, et al. Learning theories and tools for the assessment of core nursing competencies in simulation: a theoretical review. J Adv Nurs 2018 Feb;74(2):239-250. [doi: 10.1111/jan.13416] [Medline: 28815750]
23. Wong G, Greenhalgh T, Westhorp G, Pawson R. Realist methods in medical education research: what are they and what can they contribute? Med Educ 2012 Jan;46(1):89-96. [doi: 10.1111/j.1365-2923.2011.04045.x] [Medline: 22150200]
24. Wong G, Greenhalgh T, Westhorp G, Pawson R. NCBI - NIH. 2014. Development of Methodological Guidance, Publication Standards and Training Materials for Realist and Meta-narrative Reviews: The Rameses (Realist and Meta-Narrative Evidence Syntheses – Evolving Standards) Project URL: https://www.ncbi.nlm.nih.gov/pubmed/25642521 [accessed 2020-02-28]
25. Pawson R, Greenhalgh T, Harvey G, Walshe K. Semantic Scholar. 2004. Realist Synthesis: An Introduction URL: https://pdfs.semanticscholar.org/4351/46e6e6617491ff1c4b32b76e0a534c86d6c7.pdf?_ga=2.132281806.1395435436.1582877966-1679671381.1567599385 [accessed 2020-02-28]
26. Varsi C, Nes LS, Kristjansdottir OB, Kelders SM, Stenberg U, Zangi HA, et al. Implementation strategies to enhance the implementation of eHealth programs for patients with chronic illnesses: realist systematic review. J Med Internet Res 2019 Sep 27;21(9):e14255 [FREE Full text] [doi: 10.2196/14255] [Medline: 31573934]
27. Wong G, Greenhalgh T, Pawson R. Internet-based medical education: a realist review of what works, for whom and in what circumstances. BMC Med Educ 2010 Feb 2;10:12 [FREE Full text] [doi: 10.1186/1472-6920-10-12] [Medline: 20122253]
28. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]
29. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005 Jul;10(Suppl 1):21-34. [doi: 10.1258/1355819054308530] [Medline: 16053581]
30. Jeffries PR. A framework for designing, implementing, and evaluating simulations used as teaching strategies in nursing. Nurs Educ Perspect 2005;26(2):96-103. [Medline: 15921126]
31. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
32. Booth A, Sutton A, Papaioannou D. Systematic Approaches To A Successful Literature Review. Los Angeles, CA: Sage Publications Ltd; 2012.

33. Lopreiato JO, Society for Simulation in Healthcare. Agency for Healthcare Research and Quality. Healthcare Simulation Dictionary URL: https://www.ahrq.gov/sites/default/files/publications/files/sim-dictionary.pdf [accessed 2020-02-28]
34. Schoenherr JR, Hamstra SJ. Beyond fidelity: deconstructing the seductive simplicity of fidelity in simulator-based education in the health care professions. Simul Healthc 2017 Apr;12(2):117-123. [doi: 10.1097/SIH.0000000000000226] [Medline: 28704289]
35. Cant RP, Cooper SJ. The value of simulation-based learning in pre-licensure nurse education: a state-of-the-art review and meta-analysis. Nurse Educ Pract 2017 Nov;27:45-62. [doi: 10.1016/j.nepr.2017.08.012] [Medline: 28843948]
36. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs: The Four Levels. Third Edition. San Francisco: Barrett - Koehler Publishers Inc; 2006.
37. Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, et al. Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Med Res Methodol 2006 Jul 26;6:35 [FREE Full text] [doi: 10.1186/1471-2288-6-35] [Medline: 16872487]
38. Hong Q, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. mixedmethodsappraisaltoolpublic. Mixed Methods Appraisal Tool URL: http://mixedmethodsappraisaltoolpublic.pbworks.com/ [accessed 2019-11-30]
39. Wong G. The internet in medical education: a worked example of a realist review. In: Hannes K, Lockwood C, editors. Synthesizing Qualitative Research: Choosing the Right Approach. Hoboken, New Jersey: John Wiley & Sons Ltd; 2012:83-112.
40. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. J Health Serv Res Policy 2005 Jan;10(1):45-53. [doi: 10.1177/135581960501000110] [Medline: 15667704]
41. Delbecq AL, van de Ven A, Gustafson DH. Group Techniques for Program Planning: A Guide to Nominal Groups and Delphi Process. Glenview, Illinois: Scott, Foresman and Company; 1975.
42. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000 Oct;32(4):1008-1015. [Medline: 11095242]
43. Powell C. The Delphi technique: myths and realities. J Adv Nurs 2003 Feb;41(4):376-382. [doi: 10.1046/j.1365-2648.2003.02537.x] [Medline: 12581103]
44. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. Inf Manag 2004;42(1):15-29. [doi: 10.1016/j.im.2003.11.002]
45. Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: realist syntheses. J Adv Nurs 2013 May;69(5):1005-1022. [doi: 10.1111/jan.12095] [Medline: 23356726]
46. Zachariadis M, Scott S, Barrett M. Methodological implications of critical realism for mixed-methods research. Manag Inf Syst Q 2013;37(3):855-879. [doi: 10.25300/misq/2013/37.3.09]

Abbreviations

CMO: contexts, mechanisms, and outcomes
SBL: simulation-based learning

Edited by G Eysenbach; submitted 23.09.19; peer-reviewed by S Khan, S Sullivan, T Risling; comments to author 19.11.19; revised version received 14.01.20; accepted 08.02.20; published 29.04.20.

Please cite as:
Meum TT, Slettebø Å, Fossum M
Improving the Use of Simulation in Nursing Education: Protocol for a Realist Review
JMIR Res Protoc 2020;9(4):e16363
URL: http://www.researchprotocols.org/2020/4/e16363/
doi:10.2196/16363
PMID:32347808

Protocol

The Mobile Health Multiple Lifestyle Behavior Interventions Across the Lifespan (MoBILE) Research Program: Protocol for Development, Evaluation, and Implementation

Marcus Bendtsen[1*], PhD; Preben Bendtsen[1,2*], PhD; Hanna Henriksson[1*], PhD; Pontus Henriksson[1*], PhD; Ulrika Müssener[1*], PhD; Kristin Thomas[1*], PhD; Marie Löf[1,3*], PhD

[1]Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

[2]Department of Medical Specialist, Motala, Sweden

[3]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden

[*]all authors contributed equally

Corresponding Author:
Marcus Bendtsen, PhD
Department of Health, Medicine and Caring Sciences
Linköping University
581 83 Linköping
Linköping, 58183
Sweden
Phone: 46 13281000
Email: marcus.bendtsen@liu.se

Abstract

Background: Clustering of multiple lifestyle risk behaviors has been associated with a greater risk of noncommunicable diseases and mortality than one lifestyle risk behavior or no lifestyle risk behaviors. The National Board of Health and Welfare in Sweden reported in 2018 that it is important to provide additional support to individuals with multiple lifestyle risk behaviors, as risks from these behaviors are multiplicative rather than additive. However, the same report emphasized that there is a lack of knowledge regarding interventions that support changes to unhealthy lifestyle behaviors.

Objective: The MoBILE (Mobile health Multiple lifestyle Behavior Interventions across the LifEspan) research program has brought together two Swedish research groups supported by international collaborators. Through this collaboration, we aim to design and evaluate a number of novel and tailored mobile health (mHealth) multiple lifestyle behavior interventions across the life span of different health care populations. In addition, the MoBILE research program will extend ongoing research to include mHealth interventions for migrant pregnant women and children.

Methods: Each project within the MoBILE program will focus on a specific group: pregnant women, preschool children, high school and university students, and adults in primary and clinical care. All the projects will follow the same 4 phases: requirements, development, evaluation, and implementation. During the requirements phase, implementers and end users will aid the design of content and functionality of the interventions. In the development phase, findings from the first phase will be synthesized with expert domain knowledge and theoretical constructs to create interventions tailored to the target groups. The third phase, evaluation, will comprise randomized controlled trials conducted to estimate the effects of the interventions on multiple lifestyle risk behaviors (eg, alcohol, nutrition, physical activity, and smoking). The final phase will investigate how the interventions, if found effective, can be disseminated into different health care contexts.

Results: The research program commenced in 2019, and the first results will be available in 2020. Projects involving pregnant women, preschool children, and high school and university students will be completed in the first 3 years, with the remaining projects being planned for the program's final 3 years.

Conclusions: The development of evidence-based digital tools is complex, as they should be guided by theoretical frameworks, and requires large interdisciplinary teams with competence in technology, behavioral science, and lifestyle-specific areas. Individual researchers or smaller research groups developing their own tools is not the way forward, as it means reinventing the wheel over and over again. The MoBILE research program therefore aims to join forces and learn from the past 10 years of mHealth research to maximize scientific outcomes, as well as the use of financial resources to expand the growing body of evidence for mHealth lifestyle behavior interventions.

International Registered Report Identifier (IRRID): PRR1-10.2196/14894

(JMIR Res Protoc 2020;9(4):e14894) doi:10.2196/14894

KEYWORDS

telemedicine; mHealth; eHealth; life style; randomized controlled trial; focus groups

Introduction

Background

An unhealthy diet, physical inactivity, smoking, and excessive consumption of alcohol are well-established risk factors for noncommunicable diseases (NCDs), such as cardiovascular disease, cancer, respiratory disease, and type 2 diabetes [1]. The Global Burden of Disease study from 2015 [2] reported that NCDs are responsible for 70% of the deaths globally each year, and according to the World Health Organization (WHO), the total annual number of deaths from NCDs will increase to 55 million by 2030 if the trend is not reversed.

Clustering of multiple lifestyle risk behaviors has been associated with a greater risk of NCDs and mortality compared with one lifestyle risk behavior or no lifestyle risk behaviors [3]. The National Board of Health and Welfare in Sweden reported in 2018 [4] that it is important to provide additional support to individuals with multiple lifestyle risk behaviors, as risks from multiple unhealthy lifestyle behaviors are multiplicative rather than additive. However, the same report emphasized that there is a lack of knowledge regarding interventions that support changes to unhealthy lifestyle behaviors.

WHO has published a global action plan for 2013 to 2020 to reduce NCDs, with one of the objectives being to strengthen health care systems to improve prevention and self-management of people with, or at high risk for, cardiovascular disease, cancer, chronic respiratory disease, and type 2 diabetes [5]. One of the arching principles in the plan is a life-course approach. This is essential as unhealthy lifestyle behaviors tend to be established in early childhood and adolescence, tracking into adulthood [6,7]. Therefore, an effort should be made to establish evidence-based lifestyle interventions both for those who are already experiencing the negative consequences of their unhealthy lifestyle behavior as well as those in life's early stages, before risk behaviors lead to negative consequences.

Electronic Health and Mobile Health: Large Potential for Lifestyle Interventions and Disease Management in Health Care

Globally, mobile phone subscriptions increased by 97% between 2000 and 2015, with 95% of the population (7 billion people) residing in an area with a mobile cellular network [8]. The Swedish population is described as one among the most digitally mature in the world [9], with high mobile phone ownership (97%) [10]. This digitalization has radically changed communication in areas such as the travel industry and banking sector; however, it also offers great opportunities for the future of health care.

Electronic health (eHealth) has been defined by WHO as the use of information and communication technologies for health. This broad definition encapsulates electronic patient records, patient management systems, lab systems, big data, and informatics. An additional component of eHealth is mobile health (mHealth), which has been defined as a medical or public health practice that is supported by mobile devices [11]. Mobile devices include mobile phones, patient monitoring devices, and other wireless devices.

In the past 10 to 15 years, there has been a strong increase in research investigating mHealth intervention programs for improving modifiable risk factors for chronic diseases, including diet, physical activity, weight loss, smoking cessation, and alcohol consumption reduction [12-17]. mHealth has also been used for disease management to assist, inform, guide, and treat patients with various acute and chronic diseases and disorders, such as type 2 diabetes, cardiovascular disease, and mental illness [18,19]. Some of the potential benefits of using mHealth interventions instead of more traditional face-to-face interventions or disease management programs are as follows: they can be delivered at any time or place, participants do not have to visit a clinic (or the number of visits can be reduced), the programs are interactive, and they can more easily be tailored toward specific groups.

The Mobile Health Multiple Lifestyle Behavior Interventions Across the LifEspan Research Program

mHealth interventions offer new potential for the delivery of interventions that promote healthy lifestyles and self-management of NCDs, and the interest for developing such interventions has rapidly increased. However, instead of reinventing the wheel, it is time to join forces and learn from the past 10 years of mHealth research to maximize scientific outcomes and financial resources to move this research area forward.

The MoBILE (Mobile health Multiple lifestyle Behavior Interventions across the LifEspan) research program has brought together two strong Swedish research groups supported by international collaborators: The Innovative use of mobile phones to promote physical activity and nutrition across the lifespan (IMPACT) research group and the Lifestyle Intervention Implementation Research (LiiR) group.

The IMPACT group at Karolinska Institutet and Linköping University (led by Professor Marie Löf) has expertise in nutrition, physical activity, and behavioral science [20-24]. This group has more than 10 years of experience in developing mHealth interventions, with special emphasis on pregnant women and young children [25-27]. The LiiR group at Linköping University (led by Senior Lecturer Marcus Bendtsen) has expertise in medicine, occupational therapy, health psychology, and computer science and statistics. This group

has nearly 15 years of experience in developing eHealth and mHealth interventions focusing on smoking cessation, alcohol consumption, physical activity, and positive psychology [28-41]. The group has also conducted research on how to implement such interventions into daily routine in primary health care, universities, high schools, and workplaces [30,36,41-43].

This interdisciplinary cooperation includes investigators with expertise in the big 4 lifestyle risk behaviors (ie, diet, physical activity, smoking, and alcohol), behavioral science, medicine, statistics, machine learning, and information technology. Through such collaboration, we will design and evaluate a number of novel and tailored mHealth multiple lifestyle behavior interventions across the life span of different health care populations. The research program also aims to establish an excellence research center where knowledge and expertise will be disseminated to the scientific community, health professionals, and stakeholders in the eHealth and mHealth area. The program has received a 3-year grant, extendable to 6 years, from the Swedish Research Council for Health, Working Life, and Welfare (FORTE, Dnr: 2018-01410).

The motivation behind the MoBILE research program can be described by using 3 major research gaps in the field. First, we need additional evidence to decide whether multiple lifestyle behavior interventions are effective. Second, to help more individuals, we need to better understand what works and for whom, with respect to different content and functionality of mHealth interventions. Third, we need to understand how interventions can be made accessible to everyone in the community, with a special emphasis on individuals with a migrant background. The following sections provide a detailed explanation of each research gap.

Research Gap 1: Are Multiple Lifestyle Behavior Interventions Through Mobile Health the Way Forward?

Since the early 2000s, there has been great interest in multiple lifestyle behavior interventions [44]. Simultaneously addressing multiple lifestyle risk behaviors in interventions may result in increased confidence in the capacity for change and enhance effectiveness. Indeed, there is some evidence supporting that interventions targeting multiple lifestyle risk behaviors at the same time may be beneficial for improving the general lifestyle of individuals [45,46]. However, two recent meta-analyses reported modest effects regarding multiple lifestyle risk behaviors in nonclinical [44] and clinical populations [47]. Various reasons were suggested, including poor implementation of the intervention. Implementation difficulties may be hard to overcome with traditional delivery modes, such as face-to-face interventions, because of the huge demand of resources such as time and staff.

Therefore, mHealth offers new potential to this research area, because of its ability to work autonomously from health care professionals and as interventions that are digital are more flexible with respect to content and delivery than face-to-face and other traditional means of content delivery. However, in a meta-analysis of multiple lifestyle risk behaviors from 2017, only 4 of the 69 randomized controlled trials utilized eHealth technology as the sole delivery mode (eg, email, SMS text message, or website) [44]. In addition, these 4 trials were limited by low power or engagement or had questionable external validity.

In summary, mHealth provides new potential to achieve changes in unhealthy lifestyle behaviors; however, to date, the potential of mHealth interventions in this area has been scarcely explored. The MoBILE research program will utilize the complementary expertise in all 4 lifestyle behaviors and mHealth, gained by combining the efforts of research groups, to design and evaluate 7 mHealth multiple lifestyle behavior interventions across the life span of different health care populations.

Research Gap 2: What Works and for Whom in Mobile Health Interventions?

Although an increasing number of mHealth interventions are being developed for prevention and management of chronic diseases, there is a knowledge gap on how to best develop and implement such interventions, taking both patients' and health care staffs' views into consideration [48,49]. For instance, patients might prefer visiting the health care clinic in person rather than monitoring their health at home; on the other hand, the staff might expect a reduction in their workload by patients using digital tools, which might not always be the case [50]. Thus, a key area is how to motivate both patients and staff to engage in the effective use of mHealth interventions [51].

Therefore, it is important to apply a user-centered approach when developing digital interventions; however, the consideration for the implementers' expectations and readiness to implement the suggested interventions must also be taken into account [52]. The MoBILE research program will address this research gap by including both end users and implementers in the development of 7 mHealth lifestyle behavior interventions, taking into account their respective requirements and expectations.

Research Gap 3: How Do We Make Mobile Health Interventions Accessible for All?

Low socioeconomic status and migrant background are associated with inferior health; however, mobile phones are commonly accessible, irrespective of the socioeconomic status [53]. In addition, mHealth solutions can offer flexibility in terms of content, level of information (ie, advanced text, easy-to-read text, or pictures), and languages. Thus, mHealth also offers the potential to reach groups that are hard to reach with traditional face-to-face interventions in health care. However, the use of mHealth in interventions to promote healthy lifestyle behaviors among socially disadvantaged groups and migrant populations is sparse but growing rapidly. Mostly, pilot studies have been conducted thus far; however, several studies have reported promising results [54,55].

Key target groups in the Swedish context include pregnant women and young children. Approximately 25% of the women attending maternity clinics are born outside of Sweden. Thus, mHealth tools to promote a healthy lifestyle among pregnant women and their children should also be accessible to migrant populations. Consequently, a key area for research in the mHealth area is to tailor content and features to these populations.

The MoBILE research program will extend ongoing research to also include mHealth interventions for migrant pregnant women and children, as well as build on this work for other populations in the program.

Specific Aims

The overall purpose of the MoBILE research program is to design, evaluate, and implement 7 state-of-the-art mHealth multiple lifestyle behavior interventions, which can be promoted by health care professionals. These are self-management interventions that aim to support a healthy lifestyle across the life span of different populations. Specifically, the program will be built around the following aims:

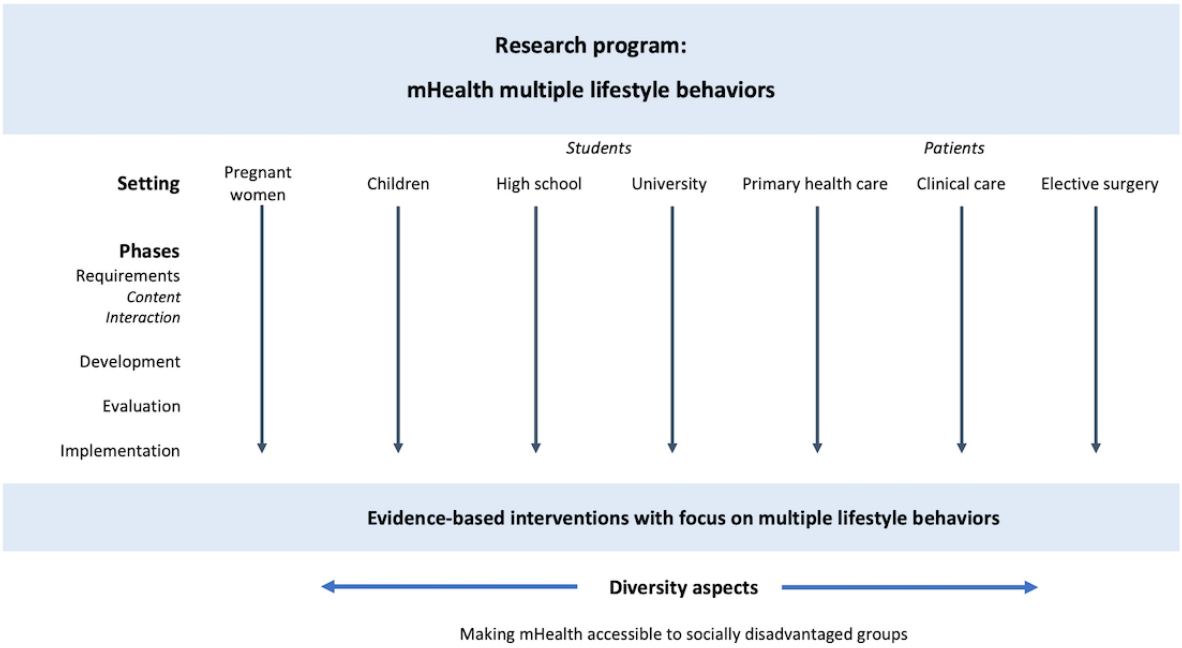
1. To investigate user requirements (patients and health care providers) for 7 mHealth multiple lifestyle behavior interventions in terms of technology and content in different health care populations (research gaps 2 and 3).
2. To assess the effectiveness of 7 mHealth multiple lifestyle behavior interventions in different health care populations (research gap 1).
3. To set a standard for how predictive, rather than explanatory, statistical modelling can be used for investigating who benefits from which multiple lifestyle behavior intervention in different health care populations (research gap 1).
4. To assess, using causal inference, the mediating effects of a number of mHealth multiple lifestyle behavior interventions through psychological factors, such as self-efficacy and motivation, in different health care populations (research gap 1).
5. To evaluate the extent to which the effectiveness, engagement, and user satisfaction for the specific mHealth tools differ among end users with different socioeconomic status and migrant backgrounds in different health care populations (research gap 3).
6. To tailor mHealth multiple lifestyle behavior interventions to groups with a migrant background in 2 health care populations (ie, pregnant women and preschool children) and to build on the knowledge gained to modify the other interventions in the program when relevant (research gap 3).
7. To implement the aforementioned mHealth multiple lifestyle behavior interventions that are deemed effective and to evaluate how well they are adopted by routine health care (research gaps 1 and 2).

Methods

Overview

A schematic presentation of the MoBILE research program is provided in Figure 1. It covers both the primary care setting as well as the specialized clinical setting and includes 7 mHealth multiple lifestyle behavior intervention projects. Each project will focus across the life span of a specific group: pregnant women, preschool children, high school and university students, and adults in primary and clinical care. All the projects will follow the same 4 phases: requirements, development, evaluation, and implementation. The rest of this section is laid out as follows: first, we introduce each of the 4 phases, which is common for all projects; thereafter, we briefly discuss the details of each project. Note that in preparation for each project, trial registration and protocols will be made available with full details before trial commencement, including recruitment and statistical analysis plans.

Figure 1. Schematic presentation of the Mobile health Multiple lifestyle Behavior Interventions across the LifEspan research program. mHealth: mobile health.



The Phases of the Mobile Health Multiple Lifestyle Behavior Interventions Across the Lifespan Research Program

Phase 1: Requirements

All proposed projects include 2 types of users: end users and implementers. End users are those individuals who use the intervention with the goal of changing their unhealthy lifestyle behaviors. Implementers are health care professionals who will administer and recommend the intervention to the end users. We will gather requirements from both types of users, specifically with respect to the content of the interventions and how the human-computer interaction should be designed.

Content Requirements

The content for each intervention will initially be based on the current best practice, gathered from scientific literature and experts. End users' and implementers' perspectives on the content of the interventions will then be explored through either focus group or individual interviews. An iterative process will be employed, which will allow content to be added, removed, or reworded in between interviews. Apart from the content itself, understandability and usability in terms of complexity of language and tone will be evaluated and adjusted for each project.

Interaction Requirements

Human-computer interaction and usability will be investigated through both heuristic evaluation [56] by interaction experts and usability tests with end users [57]. Several mock designs of the interventions will be produced, which will be analyzed by interaction experts according to a set of predefined heuristics (eg, a consistent design and error prevention). The mock designs will then be refined according to expert feedback. End users will then be asked to explore the mock designs without any guidance. While doing so, we will observe whether they are able to complete the goals that are given to them, for example, entering their daily alcohol consumption or finding information about how to deal with nicotine cravings. Refinement of the mock designs to remove obstacles that hinder the end users from completing the goals will be done in an iterative manner.

Phase 2: Development

The final content of the interventions will be decided by synthesizing the requirements gathered from phase 1. Behavior change technique analyses [58] will be conducted to elucidate how content connects to behavior change theory, for example, how motivation and self-regulation are addressed and supported by the content. The design of the interactive components that deliver the content of the interventions will be decided from the human-computer studies from phase 1.

The interventions will then be programmed and tested. Pilot studies will be conducted with end users to resolve technical issues and investigate how the final interventions are perceived. A questionnaire will be sent to pilot participants, with both fixed- and open-response options. Any major issues identified will be addressed before moving on to phase 3.

Phase 3: Evaluation

Randomized controlled trials (RCTs) will be conducted to explore the effects of the interventions in their respective target groups. The trial for each intervention aims to achieve the following:

1. Estimate the total effect of the intervention on individual lifestyle risk behaviors compared with a control setting.
2. Detect interactions among unhealthy lifestyle behavior changes (eg, those who stop smoking may also reduce their alcohol consumption).
3. Estimate to which degree the total effect is mediated through psychological factors (eg, self-efficacy and motivation).
4. Use baseline characteristics to predict who will benefit from the intervention.

For aims 1 and 2, regression models with appropriate distribution properties for the outcome measure will be used to estimate both the total effect of the intervention group compared with the control group on one lifestyle risk behavior and interaction effects among lifestyle risk behaviors. Regression coefficients' significance will be assessed using null hypothesis significance testing and exploring Bayesian posterior distributions [59]. For aim 3, we will use causal inference to assess how interventional effects are mediated through sociopsychological factors [60]. For aim 4, we will use machine learning models that aim to predict whether a certain individual, given the individual's baseline characteristics, will benefit from the intervention. This will allow us to identify subgroups that are not benefitting from the intervention, for which more research is required.

Outcome and Study Parameters

The primary outcome measures used in the RCTs will concern participants' general lifestyle, in terms of alcohol consumption, physical activity, smoking, and diet. The method through which these outcomes are measured will differ slightly among projects with respect to feasibility. Secondary outcome measures will be project specific, for example, measuring complications after surgery and body weight. We will also measure sociopsychological characteristics such as self-efficacy and motivation as mediating factors.

Power Considerations and Study Sizes

In general, mHealth interventions have an effect on lifestyle corresponding to a standardized difference between outcomes in the intervention and control groups ranging from 0.20 to 0.35 (Cohen *d*) [17,61-67]. To identify the statistically significant differences of these magnitudes using two-sided *t* tests with 80% power at a 5% significance level, it will be necessary to recruit between approximately 150 and 400 participants per arm (separate power calculations will be conducted for each project).

Phase 4: Implementation

If effective, interventions will be implemented in routine practice, and the adoption of the interventions will be assessed by the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework, which includes 5 elements [68]:

- *Reach*: The proportion of representative end users who engage with the intervention.

- *Efficacy*: The impact of the intervention on important outcomes.
- *Adoption*: The proportion of implementers who are willing to administer or recommend the intervention.
- *Implementation*: The consistency of delivery of the intervention in terms of implementers' responsibilities and end users' use of the intervention.
- *Maintenance*: To which extent the intervention becomes institutionalized and a part of routine practice.

Efficacy and long-term effects as part of maintenance will already have been assessed in phase 3. An initial estimation of reach will also have been done by comparing the number of invited participants, who are willing to partake, with the RCT.

To further investigate the reach, along with adoption, implementation, and maintenance, each project will assess the aspects of implementation in its respective health care unit. Implementers will be offered full access to administer the intervention. Adoption will be assessed through interviews, examining implementers' willingness to administer and recommend the intervention. Implementation will be assessed by engagement over time by implementers as well as engagement of end users with the intervention itself. Finally, maintenance will be measured by interest from implementers to continue using the intervention as part of their routine care as well as through a questionnaire created around the normalization process theory coding framework on eHealth implementations [69].

Project-Specific Details

The 7 main projects included in the MoBILE research program are described below.

Project 1: Pregnant Women

One project will extend ongoing research to develop mHealth tools to support a healthy lifestyle in pregnant women. We are currently evaluating the effect of an mHealth intervention (HealthyMoms app) on weight gain, diet, and physical activity during pregnancy in women who can speak and read Swedish well enough to benefit from the intervention [21]. We will extend this work by creating an mHealth solution (Healthy Migrant Moms) that is suitable for migrant women who do not understand Swedish well enough to use the original HealthyMoms app. Participants for the focus groups and panel evaluations will include both midwives and pregnant women recruited through the maternity clinic, Kvinnohälsan, Linköping. For the pregnant women, a series of focus group and individual interviews with different geographical and cultural backgrounds will be conducted with certified interpreters.

The subsequent RCT will follow a two-arm parallel-group design, for which participants will be recruited from the same maternity clinic. Both groups will receive standard antenatal care. However, the intervention group will also receive the novel intervention (Healthy Migrant Moms).

Project 2: Preschool Children

One project will implement a novel mHealth intervention [27], integrated into routine care, targeting parents of preschool children to promote the following main healthy behaviors in preschoolers: a healthy diet and physical activity. Participants for the focus groups and panel evaluations (ie, parents and nurses) will be recruited at the child health care in Östergötland. Part of this project includes translation and modification to also target priority populations such as children with a migrant background (approximately 25% of all children). Thus, focus group and individual interviews with parents will be conducted with both Swedish-speaking parents and with parents of different geographical and cultural backgrounds (using certified interpreters).

The subsequent RCT will follow a two-arm parallel-group design, for which participants will be recruited from child health care in Östergötland. Both groups will receive standard antenatal care. However, the intervention group will also receive the novel intervention.

Projects 3 and 4: High School and University Students

Two projects will develop mHealth interventions aimed at promoting improved healthy lifestyle behaviors among high school and university students. The interventions will be tailored to the two student groups based on input from the focus groups and panel evaluations.

Participants for the focus groups and interaction, as well as usability evaluations, will be recruited from high schools and student health care centers in Östergötland and from Linköping and Luleå University. A series of focus group and individual interviews with participants with different cultural backgrounds, of different age groups, and of different gender will be conducted.

An RCT for each student group will be conducted in collaboration with the student health care centers. The RCTs will follow a two-arm parallel-group design with an intervention group and a waiting-list control group. Participants will be recruited from high schools and universities located across Sweden.

Project 5: Individuals Seeking Help at Primary Health Care Centers

Healthy lifestyle promotion has been difficult to implement in routine primary health care. This project aims to develop an mHealth intervention targeting individuals seeking help at primary health care centers, for whom a lifestyle change is part of their treatment.

The focus groups and panel evaluations will comprise patient and health professional representatives from primary health care centers. We will conduct focus group and individual interviews with patients and health care professionals.

Participants for the RCT will be recruited from primary health care centers in Sweden. Health care professionals will recruit patients for whom they have determined that a lifestyle change should be part of their treatment. The trial will have a two-arm parallel-group design, in which the control group will be put on a waiting list.

Project 6: Patients in Clinical Care

One project will focus on self-care disease management and lifestyle risk behavior change among patients. The aim of the

project is to enable the patients in clinical care to change their unhealthy lifestyle behaviors, monitor symptoms and signs from their condition, or learn more about self-management.

Focus groups and panel evaluations will be conducted, comprising both patients and professionals. The content of the intervention will therefore be tailored to each disease group.

An RCT will be conducted to assess the effect of the tailored intervention. Participants will be recruited from medical wards in Sweden. The trial will follow a two-arm parallel-group design, in which the control group will be put on a waiting list.

Project 7: Patients With Elective Surgery

This project will extend ongoing research on mHealth support for patients who need to quit smoking before surgery. The project aims to add support to the existing intervention for reducing alcohol consumption, increasing physical activity, and promoting a healthy diet.

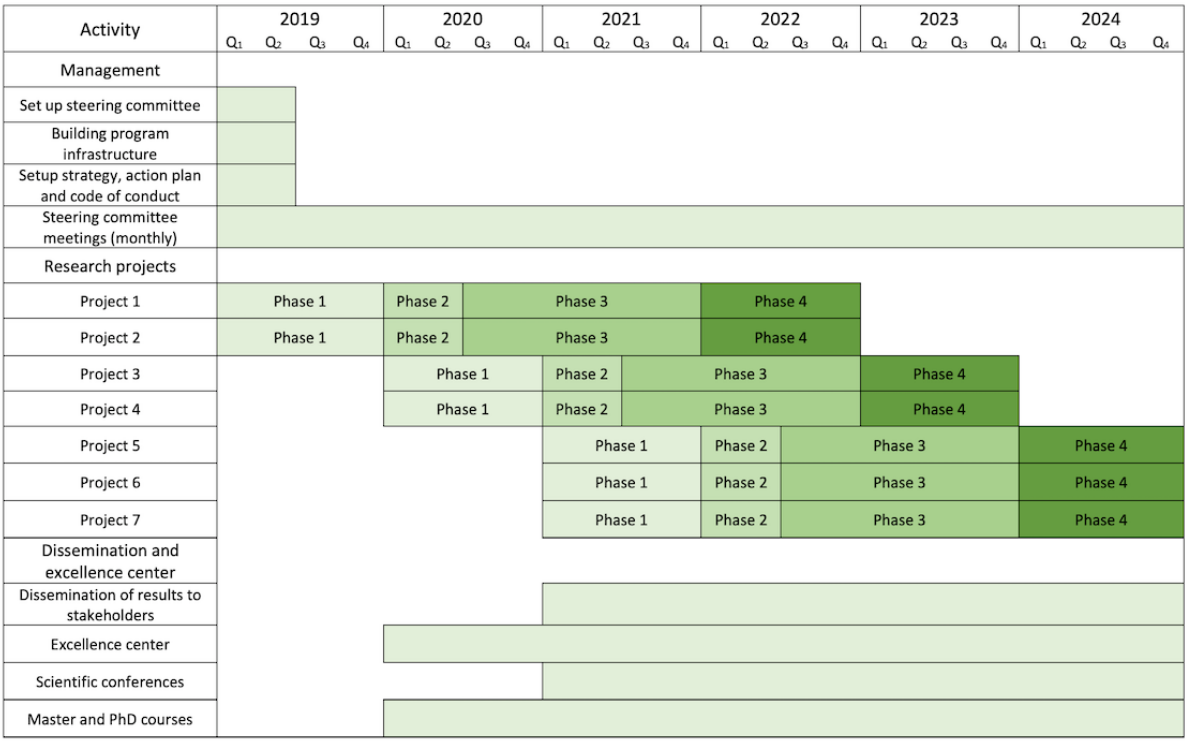
Participants for the focus groups and panel evaluations will be solicited from surgical departments in Sweden. Focus group and individual interviews will be conducted with patients and health care professionals.

Surgical departments across Sweden will be the base for recruitment for the RCT, which will have a 2-arm single-blind design. The control setting will be treatment as usual, depending on the local routines in each department. Staff at the surgical departments will inform all patients with elective surgery about the study, and consenting participants will be able to sign up on their own by using their mobile phone.

Results

The research program was funded in 2018 (FORTE, Dnr: 2018-01410, ML). The research program commenced in 2019 and is progressing according to the time plan presented in Figure 2. Briefly, the program will start with a short period where we will set up its infrastructure. Thereafter, the 7 projects (following the same 4 phases) will be executed. Projects involving pregnant women, preschool children, and high school and university students will be completed in the first 3 years, with the remaining projects being planned for the program's final 3 years. Additional activities within the research program will include the following: dissemination of the results to the stakeholder, organizing scientific workshops and conferences, and organizing master's and PhD courses on mHealth and lifestyle risk behavior topics. The first results will be available in 2020.

Figure 2. Time plan for the Mobile health Multiple lifestyle Behavior Interventions across the LifEspan research program. Phase 1: investigation of requirements by end users and implementers; phase 2: development of mobile health intervention; phase 3: randomized controlled mobile health trial; phase 4: implementation into practice.



Discussion

Joining Resources

Currently, many public health and clinical researchers are interested in developing and disseminating mHealth tools to promote a healthy lifestyle. This is a positive development; however, it also raises some concerns regarding research and evidence gaps in the field. There is still much that we do not know when it comes to the effectiveness and dissemination of mHealth interventions. Before we can launch national campaigns promoting mHealth tools, we should invest time and resources in covering the knowledge gaps that we are currently facing.

The development of evidence-based digital tools is complex, as they should be guided by theoretical frameworks, and requires

large interdisciplinary teams with competence in technology, behavioral science, and subject-specific areas (eg, nutrition and smoking cessation). Individual researchers or smaller research groups developing their own tools is not the way forward, as it means reinventing the wheel over and over again. The MoBILE research program therefore aims to join forces and learn from the past 10 years of mHealth research to maximize scientific outcomes, as well as the use of financial resources, to expand the growing body of evidence for mHealth lifestyle behavior interventions.

Clinical Relevance and Limitations

The projects included in the research program aim to address multiple unhealthy lifestyle behaviors across the life span of different health care populations; thus, the potential health benefits from the program as a whole are wide reaching. For instance, smoking cessation among adolescents will reduce the burden of disease in the future, whereas reduced alcohol consumption can lessen the immediate risk of harm among those drinking and others in their proximity. In addition, as the MoBILE research program focuses on multiple lifestyle risk behaviors and as risks increase nonlinearly with the number of unhealthy lifestyle behaviors an individual presents, the health benefits from the interventions may be superior to previous interventions focusing on one lifestyle risk behavior. This is particularly important as data indicate that individuals often have multiple unhealthy lifestyle behaviors [70].

The MoBILE research program focuses on identifying at-risk individuals from specific health care contexts, such as maternity clinics or primary health care services. This limits the reach of the included interventions and excludes a wider audience not captured within the defined contexts of the program. From a public health perspective, we wish to engage with as many at-risk individuals as possible; thus, our intention is to learn about mHealth multiple lifestyle behavior interventions within the health care context and then expand our research to include the general population.

Acknowledgments

The MoBILE program is funded by the Swedish Research Council for Health, Working Life, and Welfare (FORTE, Dnr: 2018-01410, ML).

Conflicts of Interest

MB and PB own a private company that develops and distributes lifestyle behavior interventions to be used in health care settings. The company has no connection to the projects in this research program. There are no for-profit motivations embedded in the research program. All other authors declare no conflict of interest.

Multimedia Appendix 1
Grant proposal reviewer report from Forte - Forskningsrådet för hälsa, arbetsliv och välfärd.
[PDF File (Adobe PDF File), 84 KB - resprot_v9i4e14894_app1.pdf]

References

1. World Health Organization. 2018 Jun 1. Noncommunicable Diseases: Key Facts URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2020-02-12]
2. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 8;388(10053):1659-1724 [FREE Full text] [doi: 10.1016/S0140-6736(16)31679-8] [Medline: 27733284]
3. Myint PK, Luben RN, Wareham NJ, Bingham SA, Khaw K. Combined effect of health behaviours and risk of first ever stroke in 20,040 men and women over 11 years' follow-up in Norfolk cohort of European Prospective Investigation of Cancer (EPIC Norfolk): prospective population study. Br Med J 2009 Feb 19;338:b349 [FREE Full text] [doi: 10.1136/bmj.b349] [Medline: 19228771]
4. Socialstyrelsen. National Guidelines For Prevention And Treatment In Unhealthy Living Habits. Sweden: Socialstyrelsen; 2018.
5. World Health Organization. Global Action Report. Global Action Report 2013.
6. Kaikkonen JE, Mikkilä V, Magnussen CG, Juonala M, Viikari JS, Raitakari OT. Does childhood nutrition influence adult cardiovascular disease risk?--insights from the Young Finns Study. Ann Med 2013 Mar;45(2):120-128. [doi: 10.3109/07853890.2012.671537] [Medline: 22494087]
7. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. Obes Rev 2016 Feb;17(2):95-107. [doi: 10.1111/obr.12334] [Medline: 26696565]
8. -. [Facts and figures]. Voen Med Zh 1977 May(5):18. [Medline: 406739]
9. Ministry of Health and Social Affairs. The Government of Sweden. Stockholm: Ministry of Health and Social Affairs; 2016. Vision for eHealth 2025 URL: https://www.government.se/4a3e02/contentassets/b0fd09051c6c4af59c8e33a3e71fff24/vision-for-ehealth-2025.pdf [accessed 2020-02-12]
10. Statistics S. Peoples Use of Computers and the Internet. Sweden: Statistics Sweden; 2016.

11. Ryu S. Book Review: mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). Healthc Inform Res 2012;18(3):231-250. [doi: 10.4258/hir.2012.18.3.231]
12. Mateo GF, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
13. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Apr;18(4):450-459. [doi: 10.1111/obr.12492] [Medline: 28187246]
14. O Rourke L, Humphris G, Baldacchino A. Electronic communication based interventions for hazardous young drinkers: a systematic review. Neurosci Biobehav Rev 2016 Sep;68:880-890. [doi: 10.1016/j.neubiorev.2016.07.021] [Medline: 27453570]
15. Smedslund G, Wollscheid S, Fang L, Nilsen W, Steiro A, Larun L. Effects of early, computerized brief interventions on risky alcohol use and risky cannabis use among young people. Campbell Syst Rev 2017;13(1):1-192. [doi: 10.4073/csr.2017.6]
16. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
17. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016 May 20;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
18. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
19. Gliddon E, Barnes SJ, Murray G, Michalak EE. Online and mobile technologies for self-management in bipolar disorder: a systematic review. Psychiatr Rehabil J 2017 Sep;40(3):309-319. [doi: 10.1037/prj0000270] [Medline: 28594196]
20. Ek A, Alexandrou C, Nyström CD, Direito A, Eriksson U, Hammar U, et al. The Smart City Active Mobile Phone Intervention (SCAMPI) study to promote physical activity through active transportation in healthy adults: a study protocol for a randomised controlled trial. BMC Public Health 2018 Jul 16;18(1):880 [FREE Full text] [doi: 10.1186/s12889-018-5658-4] [Medline: 30012116]
21. Henriksson P, Sandborg J, Blomberg M, Alexandrou C, Maddison R, Silfvernagel K, et al. A smartphone app to promote healthy weight gain, diet, and physical activity during pregnancy (HealthyMoms): protocol for a randomized controlled trial. JMIR Res Protoc 2019 Mar 1;8(3):e13011 [FREE Full text] [doi: 10.2196/13011] [Medline: 30821695]
22. Bonn SE, Alexandrou C, Steiner KH, Wiklander K, Östenson CG, Löf M, et al. App-technology to increase physical activity among patients with diabetes type 2 - the DiaCert-study, a randomized controlled trial. BMC Public Health 2018 Jan 10;18(1):119 [FREE Full text] [doi: 10.1186/s12889-018-5026-4] [Medline: 29316905]
23. Ek A, Sandborg J, Nyström CD, Lindqvist A, Rutberg S, Löf M. Physical Activity and Mobile Phone Apps in the Preschool Age: Perceptions of Teachers and Parents. JMIR Mhealth Uhealth 2019 Apr 17;7(4):e12512 [FREE Full text] [doi: 10.2196/12512] [Medline: 30994465]
24. Lagerros YT, Sandin S, Bexelius C, Litton J, Löf M. Estimating physical activity using a cell phone questionnaire sent by means of short message service (SMS): a randomized population-based study. Eur J Epidemiol 2012 Jul;27(7):561-566. [doi: 10.1007/s10654-012-9708-4] [Medline: 22744733]
25. Nyström CD, Forsum E, Henriksson H, Trolle-Lagerros Y, Larsson C, Maddison R, et al. A mobile phone based method to assess energy and food intake in young children: a validation study against the doubly labelled water method and 24 h dietary recalls. Nutrients 2016 Jan 15;8(1):pii: E50 [FREE Full text] [doi: 10.3390/nu8010050] [Medline: 26784226]
26. Leppänen MH, Nyström CD, Henriksson P, Pomeroy J, Ruiz JR, Ortega FB, et al. Physical activity intensity, sedentary behavior, body composition and physical fitness in 4-year-old children: results from the ministop trial. Int J Obes (Lond) 2016 Jul;40(7):1126-1133. [doi: 10.1038/ijo.2016.54] [Medline: 27087109]
27. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. Am J Clin Nutr 2017 Jun;105(6):1327-1335. [doi: 10.3945/ajcn.116.150995] [Medline: 28446496]
28. McCambridge J, Bendtsen P, Bendtsen M, Nilsen P. Alcohol email assessment and feedback study dismantling effectiveness for university students (AMADEUS-1): study protocol for a randomized controlled trial. Trials 2012 Jul 6;13:49 [FREE Full text] [doi: 10.1186/1745-6215-13-49] [Medline: 22540638]
29. Bendtsen P, McCambridge J, Bendtsen M, Karlsson N, Nilsen P. Effectiveness of a proactive mail-based alcohol Internet intervention for university students: dismantling the assessment and feedback components in a randomized controlled trial. J Med Internet Res 2012 Oct 31;14(5):e142 [FREE Full text] [doi: 10.2196/jmir.2062] [Medline: 23113955]
30. Thomas K, Müssener U, Linderoth C, Karlsson N, Bendtsen P, Bendtsen M. Effectiveness of a text messaging-based intervention targeting alcohol consumption among university students: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jun 25;6(6):e146 [FREE Full text] [doi: 10.2196/mhealth.9642] [Medline: 29941417]

31. Bendtsen M, McCambridge J. Reducing alcohol consumption among risky drinkers in the general population of sweden Using an interactive mobile health intervention: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Apr 18;8(4):e13119 [FREE Full text] [doi: 10.2196/13119] [Medline: 30998221]
32. Bendtsen M. Text messaging interventions for reducing alcohol consumption among harmful and hazardous drinkers: protocol for a systematic review and meta-analysis. JMIR Res Protoc 2019 Apr 23;8(4):e12898 [FREE Full text] [doi: 10.2196/12898] [Medline: 31012866]
33. Bendtsen M. Electronic screening for alcohol use and brief intervention by email for university students: reanalysis of findings from a randomized controlled trial using a Bayesian framework. J Med Internet Res 2019 Nov 7;21(11):e14419 [FREE Full text] [doi: 10.2196/14419] [Medline: 31697242]
34. McCambridge J, Bendtsen M, Karlsson N, White IR, Bendtsen P. Alcohol assessment & feedback by e-mail for university student hazardous and harmful drinkers: study protocol for the AMADEUS-2 randomised controlled trial. BMC Public Health 2013 Oct 10;13:949 [FREE Full text] [doi: 10.1186/1471-2458-13-949] [Medline: 24456668]
35. Bendtsen P, Bendtsen M, Karlsson N, White IR, McCambridge J. Online alcohol assessment and feedback for hazardous and harmful drinkers: findings from the AMADEUS-2 randomized controlled trial of routine practice in Swedish universities. J Med Internet Res 2015 Jul 9;17(7):e170 [FREE Full text] [doi: 10.2196/jmir.4020] [Medline: 26159179]
36. Bendtsen M, Bendtsen P. Feasibility and user perception of a fully automated push-based multiple-session alcohol intervention for university students: randomized controlled trial. JMIR Mhealth Uhealth 2014 Jun 23;2(2):e30 [FREE Full text] [doi: 10.2196/mhealth.3233] [Medline: 25098296]
37. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. SMS-based smoking cessation intervention among university students: study protocol for a randomised controlled trial (NEXit trial). Trials 2015 Apr 8;16:140 [FREE Full text] [doi: 10.1186/s13063-015-0640-2] [Medline: 25872503]
38. Thomas K, Bendtsen M, Linderoth C, Karlsson N, Bendtsen P, Müssener U. Short message service (SMS)-based intervention targeting alcohol consumption among university students: study protocol of a randomized controlled trial. Trials 2017 Apr 4;18(1):156 [FREE Full text] [doi: 10.1186/s13063-017-1898-3] [Medline: 28372563]
39. Müssener U, Bendtsen M, McCambridge J, Bendtsen P. User satisfaction with the structure and content of the NEXit intervention, a text messaging-based smoking cessation programme. BMC Public Health 2016 Nov 22;16(1):1179 [FREE Full text] [doi: 10.1186/s12889-016-3848-5] [Medline: 27876031]
40. Thomas K, Linderoth C, Bendtsen M, Bendtsen P, Müssener U. Text message-based intervention targeting alcohol consumption among university students: findings from a formative development study. JMIR Mhealth Uhealth 2016 Oct 20;4(4):e119 [FREE Full text] [doi: 10.2196/mhealth.5863] [Medline: 27765732]
41. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. Effectiveness of short message service text-based smoking cessation intervention among university students: a randomized clinical trial. JAMA Intern Med 2016 Mar;176(3):321-328 [FREE Full text] [doi: 10.1001/jamainternmed.2015.8260] [Medline: 26903176]
42. McCambridge J, Bendtsen M, Karlsson N, White IR, Nilsen P, Bendtsen P. Alcohol assessment and feedback by email for university students: main findings from a randomised controlled trial. Br J Psychiatry 2013 Nov;203(5):334-340 [FREE Full text] [doi: 10.1192/bjp.bp.113.128660] [Medline: 24072758]
43. Bendtsen P, Müssener U, Karlsson N, López-Pelayo H, Palacio-Vieira J, Colom J, et al. Implementing referral to an electronic alcohol brief advice website in primary healthcare: results from the ODHIN implementation trial. BMJ Open 2016 Jun 16;6(6):e010271 [FREE Full text] [doi: 10.1136/bmjopen-2015-010271] [Medline: 27311902]
44. Meader N, King K, Wright K, Graham HM, Petticrew M, Power C, et al. Multiple risk behavior interventions: meta-analyses of RCTs. Am J Prev Med 2017 Jul;53(1):e19-e30 [FREE Full text] [doi: 10.1016/j.amepre.2017.01.032] [Medline: 28258777]
45. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
46. de Vries H, Kremers S, Smeets T, Reubsaet A. Clustering of diet, physical activity and smoking and a general willingness to change. Psychol Health 2008;23(3):265-278. [doi: 10.1080/14768320701349107] [Medline: 25160478]
47. Alageel S, Gulliford MC, McDermott L, Wright AJ. Multiple health behaviour change interventions for primary prevention of cardiovascular disease in primary care: systematic review and meta-analysis. BMJ Open 2017 Jun 15;7(6):e015375 [FREE Full text] [doi: 10.1136/bmjopen-2016-015375] [Medline: 28619779]
48. Abroms LC, Whittaker R, Free C, van Alstyne JM, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
49. Sankaran S, Bonneux C, Dendale P, Coninx K. Bridging Patients' Needs and Caregivers' Perspectives to Tailor Information Provisioning During Cardiac Rehabilitation. In: Proceedings of the 32nd International BCS Human Computer Interaction Conference. 2018 Presented at: HCI'18; July 4-6, 2018; Belfast, United Kingdom p. 1-11.
50. Vuorinen A, Leppänen J, Kaijanranta H, Kulju M, Heliö T, van Gils M, et al. Use of home telemonitoring to support multidisciplinary care of heart failure patients in Finland: randomized controlled trial. J Med Internet Res 2014 Dec 11;16(12):e282 [FREE Full text] [doi: 10.2196/jmir.3651] [Medline: 25498992]

51. Wilhide III CC, Peeples MM, Kouyaté RC. Evidence-based mHealth chronic disease mobile app intervention design: development of a framework. JMIR Res Protoc 2016 Feb 16;5(1):e25 [FREE Full text] [doi: 10.2196/resprot.4838] [Medline: 26883135]
52. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? A qualitative study. Implement Sci 2011 Jan 19;6:6 [FREE Full text] [doi: 10.1186/1748-5908-6-6] [Medline: 21244714]
53. O'Brien OA, McCarthy M, Gibney ER, McAuliffe FM. Technology-supported dietary and lifestyle interventions in healthy pregnant women: a systematic review. Eur J Clin Nutr 2014 Jul;68(7):760-766. [doi: 10.1038/ejcn.2014.59] [Medline: 24781682]
54. Linke SE, Larsen BA, Marquez B, Mendoza-Vasconez A, Marcus BH. Adapting technological interventions to meet the needs of priority populations. Prog Cardiovasc Dis 2016;58(6):630-638. [doi: 10.1016/j.pcad.2016.03.001] [Medline: 26957186]
55. Verbiest ME, Corrigan C, Dalhousie S, Firestone R, Funaki T, Goodwin D, et al. Using codesign to develop a culturally tailored, behavior change mHealth intervention for indigenous and other priority communities: A case study in New Zealand. Transl Behav Med 2019 Jul 16;9(4):720-736. [doi: 10.1093/tbm/iby093] [Medline: 30388262]
56. Usability. Heuristic Evaluations and Expert Reviews URL: https://www.usability.gov/how-to-and-tools/methods/heuristic-evaluation.html [accessed 2020-02-07]
57. Usability. Running a Usability Test URL: https://www.usability.gov/how-to-and-tools/methods/running-usability-tests.html [accessed 2020-02-07]
58. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
59. Bendtsen M. A gentle introduction to the comparison between null hypothesis testing and Bayesian analysis: reanalysis of two randomized controlled trials. J Med Internet Res 2018 Oct 24;20(10):e10873 [FREE Full text] [doi: 10.2196/10873] [Medline: 30148453]
60. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods 2010 Dec;15(4):309-334. [doi: 10.1037/a0020761] [Medline: 20954780]
61. Riper H, Hoogendoorn A, Cuijpers P, Karyotaki E, Boumparis N, Mira A, et al. Effectiveness and treatment moderators of internet interventions for adult problem drinking: an individual patient data meta-analysis of 19 randomised controlled trials. PLoS Med 2018 Dec;15(12):e1002714 [FREE Full text] [doi: 10.1371/journal.pmed.1002714] [Medline: 30562347]
62. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev 2017 Sep 25;9:CD011479 [FREE Full text] [doi: 10.1002/14651858.CD011479.pub2] [Medline: 28944453]
63. Donoghue K, Patton R, Phillips T, Deluca P, Drummond C. The effectiveness of electronic screening and brief intervention for reducing levels of alcohol consumption: a systematic review and meta-analysis. J Med Internet Res 2014 Jun 2;16(6):e142 [FREE Full text] [doi: 10.2196/jmir.3193] [Medline: 24892426]
64. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]
65. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS text message interventions for smoking cessation: a meta-analysis. J Subst Abuse Treat 2015 Sep;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]
66. Nour M, Chen J, Allman-Farinelli M. Efficacy and external validity of electronic and mobile phone-based interventions promoting vegetable intake in young adults: systematic review and meta-analysis. J Med Internet Res 2016 Apr 8;18(4):e58 [FREE Full text] [doi: 10.2196/jmir.5082] [Medline: 27059765]
67. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med 2017 Apr;51(2):226-239. [doi: 10.1007/s12160-016-9846-0] [Medline: 27757789]
68. RE-AIM - Reach Effectiveness Adoption Implementation Maintenance. URL: http://www.re-aim.org [accessed 2020-02-07]
69. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
70. Folkhälsomyndigheten. Public Health Data and Public Health Studio URL: https://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/statistikdatabaser/folkhalsodata-och-folkhalsostudio/ [accessed 2020-02-07]

Abbreviations

eHealth: electronic health
IMPACT: innovative use of mobile phones to promote physical activity and nutrition across the lifespan
LiiR: Lifestyle Intervention Implementation Research
mHealth: mobile health

MoBILE: Mobile health Multiple lifestyle Behavior Interventions across the LifEspan
NCD: noncommunicable disease
RCT: randomized controlled trial
WHO: World Health Organization

Edited by G Eysenbach; submitted 31.05.19; peer-reviewed by E Sezgin, A Berman, S Sankaran; comments to author 03.10.19; revised version received 06.12.19; accepted 14.01.20; published 20.04.20.

Please cite as:
Bendtsen M, Bendtsen P, Henriksson H, Henriksson P, Müssener U, Thomas K, Löf M
The Mobile Health Multiple Lifestyle Behavior Interventions Across the Lifespan (MoBILE) Research Program: Protocol for Development, Evaluation, and Implementation
JMIR Res Protoc 2020;9(4):e14894
URL: http://www.researchprotocols.org/2020/4/e14894/
doi:10.2196/14894
PMID:32310147

Protocol

User Experience and Potential Health Effects of a Conversational Agent-Based Electronic Health Intervention: Protocol for an Observational Cohort Study

Marian Z M Hurmuz[1,2], MSc; Stephanie M Jansen-Kosterink[1,2], PhD; Harm op den Akker[1,2], PhD; Hermie J Hermens[1,2], Prof Dr

[1]eHealth Group, Roessingh Research and Development, Enschede, Netherlands

[2]Biomedical Signals and Systems Group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, Netherlands

Corresponding Author:
Marian Z M Hurmuz, MSc
eHealth Group
Roessingh Research and Development
PO Box 310
Enschede, 7500 AH
Netherlands
Phone: 31 880875771
Email: m.hurmuz@rrd.nl

Abstract

Background: While the average human life expectancy has increased remarkably, the length of life with chronic conditions has also increased. To limit the occurrence of chronic conditions and comorbidities, it is important to adopt a healthy lifestyle. Within the European project "Council of Coaches," a personalized coaching platform was developed that supports developing and maintaining a healthy lifestyle.

Objective: The primary aim of this study is to assess the user experience with and the use and potential health effects of a fully working Council of Coaches system implemented in a real-world setting among the target population, specifically older adults or adults with type 2 diabetes mellitus or chronic pain.

Methods: An observational cohort study with a pretest-posttest design will be conducted. The study population will be a dynamic cohort consisting of older adults, aged ≥55 years, as well as adults aged ≥18 years with type 2 diabetes mellitus or chronic pain. Each participant will interact in a fully automated manner with Council of Coaches for 5 to 9 weeks. The primary outcomes are user experience, use of the program, and potential effects (health-related factors). Secondary outcomes include demographics, applicability of the virtual coaches, and user interaction with the virtual coaches.

Results: Recruitment started in December 2019 and is conducted through mass mailing, snowball sampling, and advertisements in newspapers and social media. This study is expected to conclude in August 2020.

Conclusions: The results of this study will either confirm or reject the hypothesis that a group of virtual embodied conversational coaches can keep users engaged over several weeks of interaction and contribute to positive health outcomes.

Trial Registration: The Netherlands Trial Register: NL7911; https://www.trialregister.nl/trial/7911

International Registered Report Identifier (IRRID): PRR1-10.2196/16641

(JMIR Res Protoc 2020;9(4):e16641) doi:10.2196/16641

KEYWORDS

virtual coaching; effectiveness; user experience; evaluation protocol; older adults; adults; type 2 diabetes mellitus; chronic pain; healthy lifestyle

Introduction

As a result of socioeconomic development and progression in medicine and education, the average human life expectancy has increased significantly [1,2]. However, the aging population has also led to more older adults living with chronic diseases [2,3]. Although these diseases cannot be cured, their burden on patients can be reduced by adopting a healthy lifestyle [2,4,5]. To enable adoption of a healthy lifestyle, a deep understanding of personal motivation and the person's economic and social pressures is needed [6,7]. Based on these insights, personalized virtual coaching systems have been developed to support lifestyle changes [8]. For these systems, using multiple coaches is more effective than using a single coach because of the potential positive impact of vicarious persuasion as compared with direct persuasion (persuasion of the crowd instead of directly persuading the person) [9]. This insight has led to the introduction of the Council of Coaches (COUCH), a new concept for virtual coaching [10].

COUCH comprises a council of 5-6 virtual coaches. These coaches inform and motivate the user and discuss different topics about healthy living [10]. COUCH was developed in collaboration with end users, and the feasibility and usability of some parts of COUCH have already been tested in a lab setting (formative evaluations). The next step is to gain, through a summative evaluation, knowledge on the possible working mechanism and potential added value of this coaching system in a real-world setting among the target population [11]. As we do not want to interfere with the ongoing development of COUCH, we decided to develop a mature and simplified version of COUCH ready for testing in a real-world setting. This paper outlines the study protocol for this first test in the real world, which aims to evaluate the user experience with and the use and potential health effects of a fully working COUCH system implemented in a real-world setting among the target population.

Methods

Trial Design

This study protocol strictly follows the CONSORT-eHEALTH checklist [12] for the introduction and methods sections. This study is an observational cohort study with a pretest-posttest design. It is explorative and evaluative. The participants will be included for at least 5 weeks and up to a maximum of 9 weeks. The first week will consist of the preparation phase. In this phase, baseline measurements will be collected (T0). The following 4 weeks will consist of the implementation phase (T1). The participants will interact with COUCH during this phase. The last 4 weeks will consist of the facultative follow-up phase (T2). Participants can choose whether they want to interact with COUCH for these additional 4 weeks.

This study will be conducted in 2 countries (the Netherlands and Scotland) and consist of 2 rounds. Each round will include 25 participants per country. During the development phase, the technology and content were tested extensively. Therefore, during this study, we do not expect technical problems. However, if participants encounter minor technical problems during the first round, these problems will be fixed. During both rounds, content will be added to various coaches.

To properly evaluate the effectiveness of technology-supported health services, such as COUCH, in the real world is challenging [13-15], and it is currently increasingly acknowledged among experts that there is an urgent need for more pragmatic study designs to adequately evaluate technology-supported health services [13-16]. Microrandomization could be an appropriate alternative study design. The microrandomized trial was introduced by Klasnja et al [17] to overcome the limitations of current experimental methods, for instance randomized controlled trials, and to supplement the use of behavioral theory to guide the development of just-in-time adaptive interventions. As we are also interested in the effectiveness of the interaction between the user and the virtual coaches, we want to assess the applicability of the virtual coaches and the users' duration of interaction with the virtual coaches of a fully working COUCH system implemented in a real-world setting among the target population. To assess the users' interaction with the virtual coaches of COUCH, the interaction with one of the primary coaches (physical activity coach) will be microrandomized: Every time the user starts a conversation with this coach, the initiative of starting the conversation will be based on microrandomization. This microrandomization consists of the following two conditions: (1) The user takes the initiative and chooses the topic of the conversation, or (2) the system takes the initiative and automatically suggests the topic of conversation.

The predefined topics include gathering information about the user, goal setting, strategy selection, learning skills, and feedback and support.

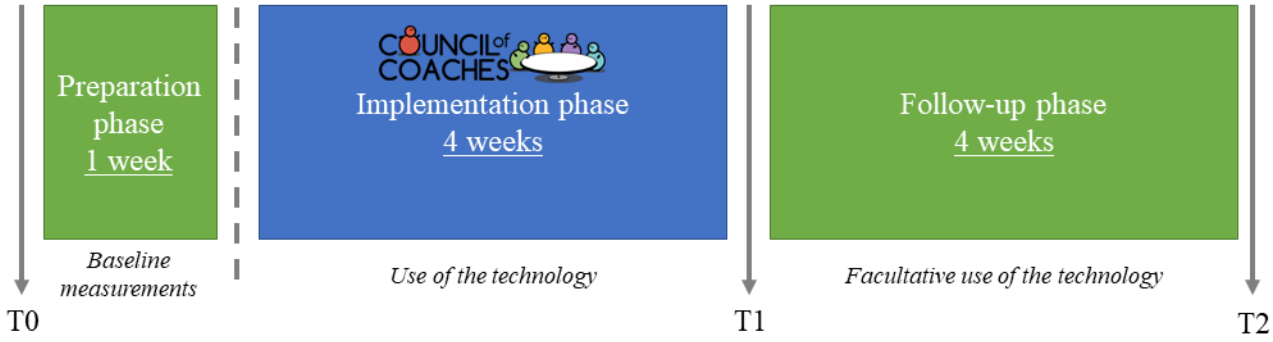
Participants

The study population will consist of older adults and adults with type 2 diabetes mellitus or chronic pain. For this study, the term older adult is defined as ≥55 years of age, and adult is defined as ≥18 years of age. A potential participant who meets any of the following criteria will be excluded from participating in this study: not able to read and speak Dutch or English, not having a Wi-Fi connection at home, not able to provide informed consent, or not able to see the smartphone or tablet screen clearly.

Eligible older adults will be recruited for the first round from December 2019 through January 2020. Participants will be recruited for the second round from March through April 2020. The first round will start in February 2020. The preparation phase (1 week) will start with an initial visit to the participant's home or an intake at the researcher's lab. During this phase, the participants do not interact with COUCH yet, but they will wear sensors for the baseline measurements, and they can keep track of their eating patterns using a food diary. The needed equipment (eg, tablet, smartphone, sensors) will be provided to the participants during this first visit, and they will receive an explanation about the operation of any equipment. Finally, participants will complete the T0 questionnaire. After this first week, the implementation phase will start. The participants will start using the COUCH system for 4 weeks. Thereafter, the second visit will take place at home or at the research location.

During this visit, an exit interview will be conducted. The participants will complete the T1 questionnaire, and they will choose whether they want to continue using COUCH for another 4 weeks (the facultative follow-up phase). If they do not want to keep using COUCH, they will return the borrowed equipment to the research staff. After the follow-up phase, all participants will complete the T2 questionnaire. The questionnaires (online or on paper) will be filled in at T0, T1, and T2 (see Figure 1).

Figure 1. Study procedures for the first and second rounds of this 9-week observational study with a pretest-posttest design.

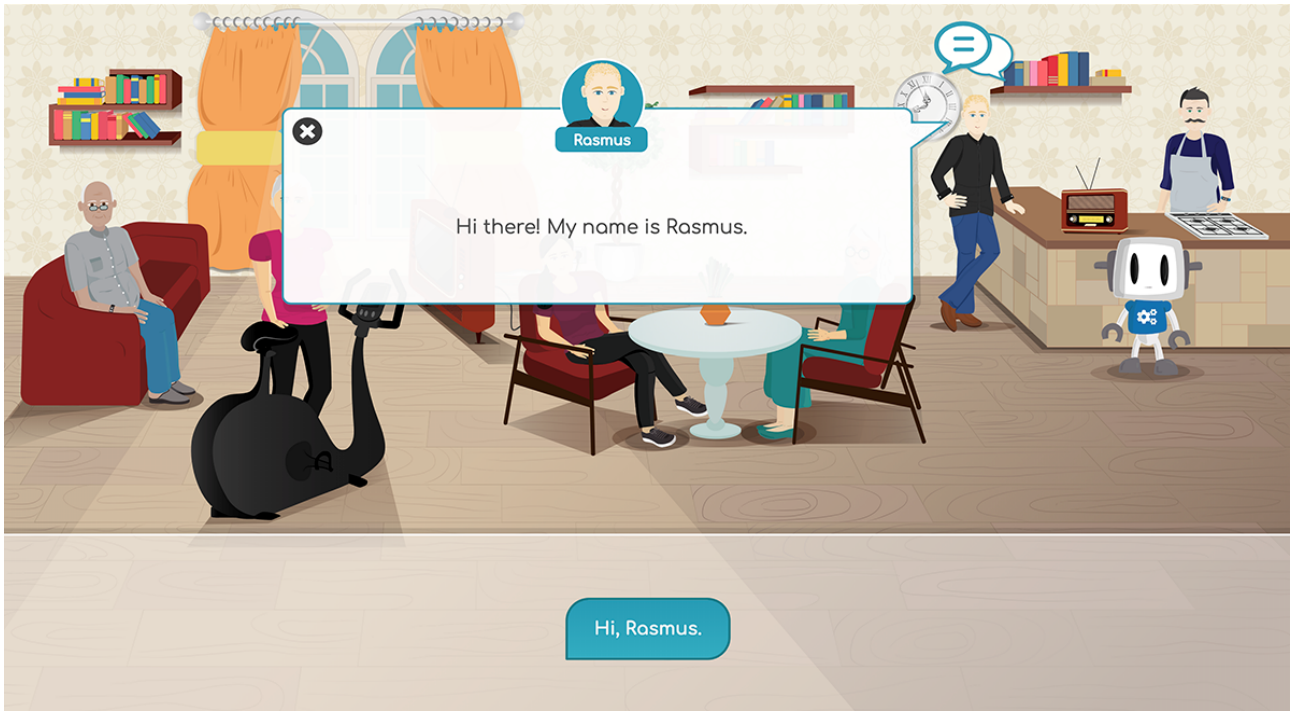


Intervention

The application is a web application, designed and built to run on tablets or computers. This technology is currently under development within the COUCH project (European Union's Horizon 2020 research and innovation program under grant agreement No. 769553). The application's main functionality is to provide a friendly and easy-to-use interface that allows users to have natural language dialogues with a group of (5-6) virtual coaches (see Figures 2 and 3). The final COUCH demonstrator will support the following virtual coaches: physical activity, nutrition, social, cognition, peer/support, chronic pain, and diabetes. Depending on the user's needs and interests, a subset of these coaches can be selected by the user (eg, in the absence of the specific conditions, the chronic pain and diabetes coaches will not be presented to the user).

Figure 2. Screenshot of the current test version of the Council of Coaches web application with the chronic pain coach, without a dialogue box (https://www.council-of-coaches.eu/beta/).

Figure 3. Screenshot of the current test version of the Council of Coaches web application with the chronic pain coach, with a dialogue box (https://www.council-of-coaches.eu/beta/).



The content provided by the virtual coaches focuses on physical fitness and nutrition to improve the users' wellbeing, and the content is based on (Dutch) health guidelines. Both the physical activity and nutrition coaches, which are the primary coaches, can assist the user in their domain in the following ways: providing information on health benefits, setting personalized goals, providing feedback and advice, reflecting on different coaching styles, and assisting with relevant sensor technology.

The secondary coaches (social, cognition, peer/support, chronic pain, and diabetes) interact with the user by providing their points of view on the main topics of activity and nutrition. For example, the social coach may suggest doing group activities outside the house when the user is discussing physical activity with his physical activity coach, while the cognitive coach can provide a memory game to do while grocery shopping for a recipe that the nutrition coach recommended. The peer/support coach is included to be "on the side of the user" and provides encouragement for the user to achieve his/her goals. The secondary coaches, except for the chronic pain and diabetes coaches, can be removed from the council by the user. The interaction with the physical activity coach will be microrandomized [17]. Every time the user starts a conversation with a primary coach, the initiative of the conversation will be taken by the system or given to the user.

The application optionally supports the use of sensor technology, in order to allow personalized feedback and coaching to the users. The physical activity coach will suggest the user wears a Fitbit watch, which is provided by the researchers to all participants, so that she may provide feedback on the user's actual activity. Similarly, the nutrition coach will ask the user to track dietary consumption through a provided smartphone app and ask the user to enter their weight information either manually or through a connected (smart) scale. Users can talk with their virtual coaches about the use of these devices, and the coaches will explain which data is collected and for what purpose and offer the ability to stop tracking data when the user feels uncomfortable about this.

All of the interactions take place in the comfort of the coaches' living room (see Figure 2) that includes elements like a radio (playing the coaches' favorite classical songs), recipe books (that Francois, the nutrition coach, can guide the user through), and a television on which to watch physical exercise examples.

During the first visit (T0), the participants will be trained by the researcher to learn how to interact with COUCH on their tablet, and they will receive a paper manual about COUCH. During the entire evaluation period, there will be a helpdesk available for the participants on working days from 9 am to 5 pm, and the participants will receive a nonpersonalized informative newsletter three times by email to inform them about the project and running evaluation.

Outcomes

In this study, we will focus mainly on user experience, potential effects on health-related factors, and the use of COUCH during the implementation and follow-up phases. Furthermore, we will examine the demographics, applicability of the virtual coaches, and user's interaction with the virtual coaches. Table 1 gives an overview of all the questionnaires that will be used during this study. All survey questions in the 3 questionnaires are listed in Multimedia Appendix 1.

Table 1. Overview of the questionnaires and when they will be used.

	T0[a]	T1[b]	T2[c]
User experience			
Technology Acceptance Model	_[d]	X	–
System Usability Scale	–	X	–
Willingness to pay	–	X	–
Potential health effects			
EQ-5D-5L[e]	X	X	X
Positive health dimensions	X	X	X
Self-Management Ability Scale – short version	X	X	X
Demographics	X	–	–
Applicability of the virtual coaches			
Rating scale	X	X	–
Working Alliance Inventory	–	X	–

[a]Baseline.

[b]After the 4-week implementation phase.

[c]After the 4-week facultative follow-up phase.

[d]Not applicable.

[e]EQ-5D-5L: 5-level EQ-5D questionnaire.

User Experience

To determine the user experience, the Technology Acceptance Model [18,19] and System Usability Scale (SUS) [20] will be used. Furthermore, an exit interview will be conducted, and the willingness to pay will be measured. In this study, user experience domains will be used as external variables. In the literature, 4 constructs are found for the user experience of electronic health (eHealth) services. The first is enjoyment. van der Heijden [21] defined perceived enjoyment of a technology as the extent to which fun can be derived from using the system as such. He used 4 questions on a 7-point semantic differentials scale to measure the following 4 items: enjoyable – disgusting, exciting – dull, pleasant – unpleasant, and interesting – boring. The second construct is aesthetics. Lavie and Tractinsky [22] developed and validated a questionnaire to measure perceived website aesthetics. In this study, only classical aesthetics will be used. The third construct is control. In their study, van Velsen et al [23] used 3 control questions from Liu [24] that measure how users perceive the controllability of websites. The fourth construct is trust in technology. This domain is also a predictor for someone's intention to use technology [23]. van Velsen et al [23] used 4 statements about trust in technology based on the study of Harrison McKnight et al [25] about the impact of consumer trust on intentions to transact with a website.

Perceived usefulness, perceived ease of use, and intention to use will also be used as constructs in this study's questionnaire. The attitude toward the technology domain will be used as a demographic variable for the secondary outcomes. Both the perceived usefulness and perceived ease of use constructs are derived from Davis [18]. In his study, a new measurement scale for perceived usefulness and perceived ease of use was developed and validated. Both constructs are important when determining the intention to use: the less effort involved in a technology, the more it will be used, and the greater someone's belief that using the technology would enhance his/her performance, the more it will be used [18,26]. Regarding the intention to use construct, van Velsen et al [23] based this construct on those of Davis et al [19] and Gefen et al [27] and expanded it with one item of their own. Based on the study by van Velsen et al [23], 3 statements were used in this study. Those 3 items were deemed the best to assess the intention to use.

The aesthetics, control, trust in technology, perceived usefulness, perceived ease of use, and intention to use constructs all use statements rated using a 7-point Likert scale, ranging from total disagreement to total agreement.

The SUS will be used to measure the usability of COUCH. Broekhuis et al [28] showed that the SUS is insufficient as a standalone tool for assessing the usability of eHealth technologies. However, another eHealth usability tool is not yet available [28]. The SUS consists of 10 statements with 5 response options that are rated using a 5-point Likert scale ranging from strongly disagree to strongly agree. The SUS score ranges from 0 (worst imaginable) to 100 (best imaginable) points [20].

Qualitative feedback from the participants will be obtained through a short semistructured exit interview at T1 (after interacting with COUCH for 4 weeks). During this interview, participants will be asked to share their ideas about COUCH. We will discuss the advantages, points for improvement, and problems experienced.

Willingness to pay will be measured by asking whether the participants are willing to pay for COUCH, and, if so, how many Euros they are willing to pay.

Potential Effect on Health-related Factors

Health effects will be measured through differences in scores within the EQ-5D-5L questionnaire, 6 domains of Positive Health, and Self-Management Ability Scale – short version (SMAS-S). The EQ-5D-5L questionnaire measures quality of life and consists of a descriptive system that includes 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) and a visual analogue scale. Each dimension has 5 levels, ranging from no problems to extreme problems. With the visual analogue scale, the participants rate their health on a vertical scale, labelled from the worst health you can imagine (0) to the best health you can imagine (100) [29].

Huber et al [30] studied how people think about health. They concluded that the concept of health no longer fits within the definition of the World Health Organization (health as complete wellbeing and absence of disease). The Institute for Positive Health created a tool to gain insight into the positive health of a person. This tool consists of 6 dimensions: bodily functions, mental wellbeing, meaningfulness, quality of life, participation, and daily functioning. Participants complete the questionnaire, resulting in a score between 0 and 10 for each dimension [30]. In our study, an adapted version will be used. Instead of completing a questionnaire consisting of 42 questions, the participants score each dimension from 0 to 10, as reported by van Velsen et al [31].

The SMAS-S is a questionnaire that measures 6 self-management abilities in older adults: taking initiative, investment behavior, variety, multifunctionality, self-efficacy, and positive frame of mind. It determines whether older adults need self-management courses [32].

Use of COUCH

The actual use will be determined using the platform's log history. This outcome measure is defined as the frequency and duration of use overall, per week, and per session.

Demographics

Demographic data collected in the pretest questionnaire include gender, age, educational level, living situation, working status, attitude toward technology, self-reported level of physical activity, health literacy [33], and motivation level to live healthy. Attitude toward using technology and motivation level to live healthy will be explained in the following paragraphs.

To determine the participant's attitude toward using technology, 4 items from Agarwal and Prasad [34] are included in the questionnaire. They developed and validated a new instrument consisting of 4 statements rated using a 7-point Likert scale, ranging from total disagreement to total agreement.

To get participants engaged in working on their health, it is important to determine their motivation to live healthy. With this information, the best suitable persuasive feature can be used in COUCH for each participant [31]. The motivation of an older adult to live healthy can be measured by a tool developed by van Velsen et al [31] based on the revised Sport Motivation Scale (SMS-II). The SMS-II was created and validated by Pelletier et al [35]. This questionnaire measures sport motivation using the Self-Determination Theory. The Self-Determination Theory distinguishes between 6 types of motivation: intrinsic motivation, extrinsic external regulation, extrinsic introjected regulation, extrinsic identified regulation, extrinsic integrated regulation, and a-motivation [36]. Those 6 types are included in the SMS-II tool. According to van Velsen et al [31], there are only 3 types of motivation in older adults to live healthy: intrinsic motivation, extrinsic external regulated, and a-motivation. They provided a set of 11 statements that will be used in our study. In our study, a fourth motivation type, dual motivation, will be included because some participants are not obviously intrinsically motivated nor externally motivated.

Applicability of the Virtual Coaches

The applicability of the virtual coaches will be measured using a rating scale and an adapted version of the Working Alliance Inventory Dutch version for use in the rehabilitation setting. This questionnaire will be completed for the 2 primary virtual coaches. This questionnaire measures how the patient feels about the therapeutic alliance: the better the therapeutic alliance, the more likely the patient will follow the treatment faithfully. Each participant will provide a score between 12 and 60: the higher the score, the more satisfied the participant is with the physical activity or nutrition coach and the more she/he trusts the coach [37].

Sample Size

Because of the explorative character of this study, no sample size calculation was conducted beforehand. To answer the objectives of this study, the goal is to include 50 participants per country. So, in each round, 25 participants will be included per country. In our experience, participants are very enthusiastic to participate in this kind of evaluation with new technology before starting the study, but we expect that around 50% of the participants will drop out before the end of the implementation phase.

Statistical Analysis

Statistical analyses will be performed using SPSS, version 19 for Windows (IBM Corp, Armonk, NY). For all analyses, the CIs will be set at 95%. Descriptive statistics, such as frequency, mean, SD, and percentages, will be used to describe demographics, user experience, actual use, and the applicability of the coaches.

The outcomes from the EQ-5D-5L, Positive Health questionnaire, and SMAS-S will be investigated using a mixed-model analysis for repeated measures to obtain the effect of using COUCH on the different measurements. The fixed factor will be the measurement time point (T0, T1, or T2). Post hoc comparisons will be conducted when required, and Sidak adjustments will be used to correct for multiple tests.

To assess the users' interaction with the virtual coaches, the duration of the interaction (in seconds) and the number of dialogue steps with the coach will be used. With this analysis, we want to assess the effect of the conversation with the virtual

coaches. To discover changes and possible trends, the duration of the interaction and number of dialogue steps will be analyzed for the two conditions. When the data follow a normal distribution, the outcome will be investigated using a paired *t* test; when the data are not normally distributed, a Wilcoxon signed-rank test will be performed.

Ethics and Informed Consent

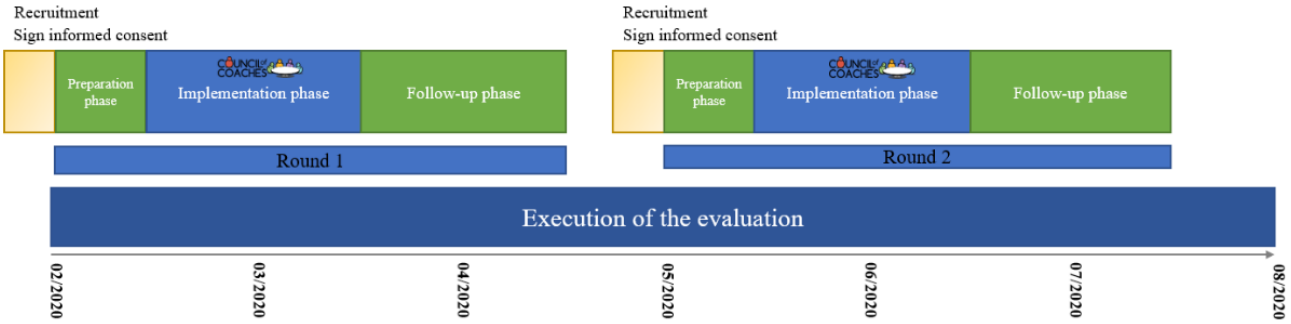
This study will be conducted according to the principles of the Declaration of Helsinki (64th WMA General Assembly, Fortaleza, Brazil, October 2013) and in accordance with the Medical Research Involving Human Subjects Act (Dutch law: Wet medisch-wetenschappelijk onderzoek met mensen). According to this law, this study does not require formal medical ethical approval in the Netherlands. This has been checked by the CMO Arnhem-Nijmegen (file number: 2019-5555). Each participant will give his/her informed consent on paper. See Multimedia Appendix 2 for the informed consent form.

Results

Recruitment of participants will take place twice. The first round of recruitment occurred from December 1, 2019 to January 30, 2020 in the Netherlands, during which time we recruited 26 participants. The first round of recruitment is still ongoing in Scotland. The second round of recruitment will occur from March 1, 2020 to April 30, 2020. For each round, we will recruit 25 participants per country. Participants are recruited through a mass mailing to older adults, snowball sampling, and advertisements in local newspapers and social media. Participants contact the principal investigator to sign up for participation. The principal investigator sends interested individuals an information letter via email and checks the inclusion and exclusion criteria. If a participant is eligible and still wants to be enrolled in the study, the first visit is planned by the principal investigator, and the study starts.

The first round of evaluation started on January 31, 2020. This round will last until April 15, 2020. The second round of evaluation will start in May 2020 and will last until July 2020. Figure 4 shows the planning of the evaluation. In August 2020, we plan to have the first results of this study.

Figure 4. Timeline of the study evaluation period.



Discussion

Overview

This protocol describes the final evaluation of the COUCH system. This study has the following strengths. First, the COUCH system was developed in collaboration with end users. Our expectation is that this will lead to fewer usability issues and better insight into the study's primary outcome measures. McCurdie et al [38] reported that users identify key requirements that otherwise would entirely be neglected. Second, this evaluation will take place in the participants' residence, a real-world setting, over a long period (5-9 weeks). This will provide a lot of information about how long the target group is willing to interact with a virtual coaching system and whether a virtual coaching system can lead to behavior change. Finally, the intervention will be personalized to the participants. We will start the evaluation with a 1-week baseline measurement, in which we will measure the participants' activity level and eating patterns. With this information, we can personalize the physical activity and nutrition coaches for each participant, which will improve the effectiveness of COUCH. Lentferink et al [39] showed in their scoping review that personalized content improves adherence to eHealth technologies, which subsequently will lead to a more effective eHealth service.

Limitations

However, this study also has some limitations. First, there will likely be selection bias. Participants contact the researchers to enroll in the study. We expect that these participants are already more motivated to live healthy or already live more healthily than the average older adult population and the average adult population with type 2 diabetes or chronic pain. Second, the content that will be ready at the start of the evaluation only lasts for 4 weeks. During the follow-up phase, no new content will be provided to the participants. This can influence the interaction frequency during the follow-up phase. Finally, this study will possibly have to deal with confounders, for example if users receive advice from their health care professionals or others about a healthy lifestyle. This occurs in real life. To handle this as best as possible, confounders such as these will be discussed with the users during the exit interview.

Conclusions

This study will provide insight spanning many areas to improve the COUCH system, and it will contribute to further development of the system and to a better understanding of the

value of virtual coaches for behavior change. In addition, the summative approach of this study protocol to evaluate an eHealth application in a real-world setting can be used to guide other eHealth evaluations.

Acknowledgments

This study is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 769553.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey questions.
[PDF File (Adobe PDF File), 477 KB - resprot_v9i4e16641_app1.pdf]

Multimedia Appendix 2
Informed consent form.
[PDF File (Adobe PDF File), 108 KB - resprot_v9i4e16641_app2.pdf]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1558 KB - resprot_v9i4e16641_app3.pdf]

References

1. Gulland A. Global life expectancy has risen, reports WHO. BMJ 2014 May 16;348:g3369-g3369 [FREE Full text] [doi: 10.1136/bmj.g3369]
2. Suzman R, Beard JR, Boerma T, Chatterji S. Health in an ageing world—what do we know? The Lancet 2015 Feb;385(9967):484-486 [FREE Full text] [doi: 10.1016/S0140-6736(14)61597-X]
3. van Oostrom SH, Gijsen R, Stirbu I, Korevaar JC, Schellevis FG, Picavet HSJ, et al. Time Trends in Prevalence of Chronic Diseases and Multimorbidity Not Only due to Aging: Data from General Practices and Health Surveys. PLoS One 2016 Aug 2;11(8):e0160264 [FREE Full text] [doi: 10.1371/journal.pone.0160264] [Medline: 27482903]
4. Willett W, Skerrett P, Giovannucci EL. Eat, Drink, And Be Healthy: The Harvard Medical School Guide To Healthy Eating. New York: Free Press; 2017.
5. World Health Organization. Preventing chronic diseases: a vital investment. 2005. URL: http://web.archive.org/web/20191002112801/https://www.who.int/chp/chronic_disease_report/full_report.pdf [accessed 2019-07-24]
6. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016 May 13;136:109-116 [FREE Full text] [doi: 10.1016/j.puhe.2016.03.030] [Medline: 27184821]
7. Bundy C. Changing behaviour: using motivational interviewing techniques. J R Soc Med 2004 Jan 01;97 Suppl 44:43-47 [FREE Full text] [Medline: 15239293]
8. Kulyk O, op den Akker R, Klaassen R, van Gemert-Pijnen L. Let us Get Real! An Integrated Approach for Virtual Coaching and Real-time Activity Monitoring in Lifestyle Change Support Systems. In: Proceedings of the 6th International Conference on eHealth, Telemedicine, and Social Medicine. Wilmington: IARIA; 2014 Presented at: eTELEMED; 2014 Mar 23-27; Barcelona, Spain p. 211-216 URL: https://tinyurl.com/wadvoxg
9. Kantharaju RB, De Franco D, Pease A, Pelachaud C. Is Two Better than One? Effects of Multiple Agents on User Persuasion. In: Proceedings of the 18th International Conference on Intelligent Virtual Agents. New York: ACM; 2018 Presented at: ICA2018; 2018 Nov 5-8; Sydney, Australia p. 255-262. [doi: 10.1145/3267851.3267890]
10. op den Akker H, op den Akker R, Beinema T, Banos O, Heylen D, Bedsted B, et al. Council of Coaches - A Novel Holistic Behavior Change Coaching Approach. In: Proceedings of the 4th International Conference on Information Communication Technology for Ageing Well and e-Health. Setúbal: SciTePress; 2018 Presented at: ICT4AWE; 2018 Mar 22-23; Funchal, Madeira, Portugal p. 219-226. [doi: 10.5220/0006787702190226]
11. Jansen-Kosterink S, Vollenbroek-Hutten M, Hermens H. A Renewed Framework for the Evaluation of Telemedicine. In: Proceedings of the 8th International Conference on eHealth, Telemedicine, and Social Medicine. Wilmington: IARIA; 2016 Presented at: eTELEMED; 2016 Apr 24-28; Venice, Italy.
12. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

13. Kairy D, Lehoux P, Vincent C, Visintin M. A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation. Disabil Rehabil 2009 Jul 07;31(6):427-447. [doi: 10.1080/09638280802062553] [Medline: 18720118]
14. Laplante C, Peng W. A systematic review of e-health interventions for physical activity: an analysis of study design, intervention characteristics, and outcomes. Telemed J E Health 2011 Aug 18;17(7):509-523. [doi: 10.1089/tmj.2011.0013] [Medline: 21718092]
15. Ekeland AG, Bowes A, Flottorp S. Methodologies for assessing telemedicine: A systematic review of reviews. International Journal of Medical Informatics 2012 Jan;81(1):1-11. [doi: 10.1016/j.ijmedinf.2011.10.009] [Medline: 22104370]
16. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: A systematic review of reviews. International Journal of Medical Informatics 2010 Nov;79(11):736-771 [FREE Full text] [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
17. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. Health Psychology 2015;34(Suppl):1220-1228. [doi: 10.1037/hea0000305] [Medline: 26651463]
18. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
19. Davis FD, Bagozzi RP, Warshaw PR. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 1989 Aug 1;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
20. Brooke J. SUS - A quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London: Crc Press; 1996:189-194.
21. van der Heijden H. User Acceptance of Hedonic Information Systems. MIS Quarterly 2004 Dec;28(4):695-704. [doi: 10.2307/25148660]
22. Lavie T, Tractinsky N. Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human-Computer Studies 2004 Mar;60(3):269-298. [doi: 10.1016/j.ijhcs.2003.09.002]
23. van Velsen L, van der Geest T, van de Wijngaert L, van den Berg S, Steehouder M. Personalization has a Price, Controllability is the Currency: Predictors for the Intention to use Personalized eGovernment Websites. Journal of Organizational Computing and Electronic Commerce 2015 Feb 05;25(1):76-97. [doi: 10.1080/10919392.2015.990782]
24. Liu Y. Developing a scale to measure the interactivity of websites. JAR 2003 Jun 01;43(2):207-216. [doi: 10.2501/JAR-43-2-207-216]
25. Harrison McKnight D, Choudhury V, Kacmar C. The impact of initial consumer trust on intentions to transact with a web site: a trust building model. The Journal of Strategic Information Systems 2002 Dec;11(3-4):297-323 [FREE Full text] [doi: 10.1016/s0963-8687(02)00020-3]
26. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science 2000 Feb 1;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
27. Gefen D, Karahanna E, Straub DW. Trust and TAM in Online Shopping: An Integrated Model. MIS Quarterly 2003 Mar;27(1):51-90. [doi: 10.2307/30036519]
28. Broekhuis M, van Velsen L, Hermens H. Assessing usability of eHealth technology: A comparison of usability benchmarking instruments. Int J Med Inform 2019 Aug;128:24-31. [doi: 10.1016/j.ijmedinf.2019.05.001] [Medline: 31160008]
29. van Reenen M, Janssen B. EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. 2019 Sep. URL: https://tinyurl.com/vscyjkf [accessed 2019-03-08]
30. Huber M, van Vliet M, Giezenberg M, Winkens B, Heerkens Y, Dagnelie PC, et al. Towards a 'patient-centred' operationalisation of the new dynamic concept of health: a mixed methods study. BMJ Open 2016 Jan 12;6(1):e010091 [FREE Full text] [doi: 10.1136/bmjopen-2015-010091] [Medline: 26758267]
31. van Velsen L, Broekhuis M, Jansen-Kosterink S, Op den Akker H. Tailoring Persuasive Electronic Health Strategies for Older Adults on the Basis of Personal Motivation: Web-Based Survey Study. J Med Internet Res 2019 Sep 06;21(9):e11759 [FREE Full text] [doi: 10.2196/11759] [Medline: 31493323]
32. Schuurmans H, Steverink N, Frieswijk N, Buunk BP, Slaets JPJ, Lindenberg S. How to measure self-management abilities in older people by self-report. The development of the SMAS-30. Qual Life Res 2005 Dec;14(10):2215-2228. [doi: 10.1007/s11136-005-8166-9] [Medline: 16328901]
33. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
34. Agarwal R, Prasad J. A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. Information Systems Research 1998 Jun 01;9(2):204-215. [doi: 10.1287/isre.9.2.204]
35. Pelletier LG, Rocchi MA, Vallerand RJ, Deci EL, Ryan RM. Validation of the revised sport motivation scale (SMS-II). Psychology of Sport and Exercise 2013 May;14(3):329-341. [doi: 10.1016/j.psychsport.2012.12.002]
36. Deci EL, Ryan RM. Overview of self-determination theory: an organismic dialectical perspective. In: Handbook Of Self-Determination Research. Rochester: University Of Rochester Press; 2004:3-33.

37. Paap D, Schrier E, Dijkstra PU. Development and validation of the Working Alliance Inventory Dutch version for use in rehabilitation setting. Physiother Theory Pract 2018 May 07;35(12):1292-1303. [doi: 10.1080/09593985.2018.1471112] [Medline: 29733745]
38. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;46(S2):49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
39. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key Components in eHealth Interventions Combining Self-Tracking and Persuasive eCoaching to Promote a Healthier Lifestyle: A Scoping Review. J Med Internet Res 2017 Aug 01;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]

Abbreviations

COUCH: Council of Coaches.
eHealth: electronic health.
SMAS-S: Self-Management Ability Scale – short version.
SMS-II: revised Sports Motivation Scale.
SUS: System Usability Scale.

Edited by G Eysenbach; submitted 10.10.19; peer-reviewed by R Barak Ventura, V Kukreti; comments to author 05.12.19; revised version received 23.01.20; accepted 07.02.20; published 03.04.20.

Please cite as:
Hurmuz MZM, Jansen-Kosterink SM, op den Akker H, Hermens HJ
User Experience and Potential Health Effects of a Conversational Agent-Based Electronic Health Intervention: Protocol for an Observational Cohort Study
JMIR Res Protoc 2020;9(4):e16641
URL: https://www.researchprotocols.org/2020/4/e16641
doi:10.2196/16641
PMID:32242517

Protocol

Promoting Wellness Through Mobile Health Technology in a College Student Population: Protocol Development and Pilot Study

Susanne B Haga[1], PhD; Ryan Shaw[2], PhD; Charles Kneifel[3], PhD; Sarah J Bond[4]; Geoffrey S Ginsburg[1], MD, PhD

[1]Center for Applied Genomics & Precision Medicine, Department of Medicine, Duke University School of Medicine, Durham, NC, United States

[2]School of Nursing, Duke University, Durham, NC, United States

[3]Office of Information Technology, Duke University, Durham, NC, United States

[4]Duke University, Durham, NC, United States

Corresponding Author:
Susanne B Haga, PhD
Center for Applied Genomics & Precision Medicine
Department of Medicine
Duke University School of Medicine
101 Science Drive
Box 3382
Durham, NC, 27708
United States
Phone: 1 919 684 0325
Email: susanne.haga@duke.edu

Abstract

Background: The health and well-being of college students has garnered widespread attention and concern in recent years. At the same time, the expansion and evaluation of digital technologies has grown in recent years for different target populations.

Objective: This protocol aims to describe a pilot feasibility study on wearables to assess student interest and to gather baseline data from college freshmen, for the academic year 2019 to 2020.

Methods: All full-time college freshmen residing in a single residence hall were eligible to participate. Study invitations were sent by post and email 5 weeks prior to move-in. Web-based enrollment and in-person attendance at study orientation sessions were mandatory. We provided the incoming freshmen with a wearable and study app. Wearable data and weekly survey data will be collected through the study app and analyzed. We have collected demographic, enrollment, and attrition data and the number and type of support requests from students.

Results: The planning phase of the WearDuke initiative was completed in 2018 to 2019, and the pilot study was launched in July 2019. Of the 175 students invited, 120 enrolled and 114 started the study; 107 students remained active participants till the end of the fall semester. For Apple Watch participants (the majority of study population), weekly survey completion rates ranged from 70% (74/106) to 96% (95/99).

Conclusions: Halfway through the pilot, we noticed that the initiative has been received positively by the students with minimal attrition. The short- and long-term benefits may be substantial for students, the campus, the utilization of health services, and long-term health.

International Registered Report Identifier (IRRID): DERR1-10.2196/16474

(JMIR Res Protoc 2020;9(4):e16474) doi:10.2196/16474

KEYWORDS

college students; wearables; wellness

Introduction

Background

The transition to college can be an emotionally challenging time, with new experiences, pressures, choices, and independence [1-5]. Colleges and universities have developed a wide range of resources to support student health, well-being, and academic success. Yet, while these efforts are focused on students' college years, there remain unmet needs, eg, methods that enable students to monitor health-related behaviors, an awareness of the consequences of unhealthy behaviors, and the establishment of healthy behaviors for the future. Many students have a mobile phone (over 96% of adults aged 18-29 years have reported owning one [6]). Thus, apps and tethered devices can serve as useful tools to track daily habits and health-related variables. These data and tools can help identify trends over the week or a semester and utilize reminders, sleep goals, and other types of alerts to prompt behavior change. Although tracking behavior change through tools, such as apps, has been linked to improved well-being [7,8], there is limited research on the use of wearable technologies for improving well-being, particularly among college students who are transitioning to adulthood. To this end, we are launching a scalable, hybrid population health, research, and educational initiative focused on an undergraduate student population. The goals of the initiative are to promote health awareness and engagement, establish individual healthy behaviors, promote a campus culture that fosters healthy living, and provide unique student-centered research and learning opportunities.

Our initiative in promoting student health and wellness capitalizes on 2 major disciplines: (1) population health and (2) digital technologies. Population health has become a central focus of research and an overarching goal for health systems and communities [9]. This focus on population health does not preclude individual-based interventions, rather, the two are inextricably linked. Many factors affect population health, including social determinants, the physical environment, the workplace, access to care, and behaviors. On a population level, greater effort is needed for more comprehensive assessments for risk stratification and to advance the understanding of effective interventions for particular populations.

Emerging digital or mobile health (mHealth) technologies such as smartphone apps, sensors, and wearables and connected devices can collect health-related data from individuals and populations continuously in their daily environment and analyze that data, thus creating a feedback loop to deliver in-time intervention(s) that allow people to self-manage their health and make better choices [10].Furthermore, these data can provide clinicians with a more complete picture of patient health to enable more informed and precise treatment decision making [11]. Owing to the ubiquity of smartphones, the use of mHealth technologies to measure and record health-related behaviors and clinical parameters has become increasingly convenient and useful for research [12]. These tools can increase access to research opportunities across diverse populations, collect near real-time data, and reduce costs for population health research by forgoing in-person visits for assessment and reducing study staff personnel. For example, the national *All of Us* Research Program (formerly Precision Medicine Initiative) recently selected Fitbit to pilot in up to 10,000 participants to record heart rate, physical activity, and sleep.

In addition to sensors embedded in a mobile phone, there are many connected devices that tether to smartphones that allow for the collection of health-related data, eg, glucometers, wrist-worn accelerometers, wireless scales, and portable electrocardiograms. Activity trackers are some of the most popular connected devices and include Fitbit, GENEactiv, Polar, Apple Watch, Vivosmart (Garmin), and Verily Study Watch, among others. Many of these devices also track sleep by largely relying on actigraphy and utilize an accelerometer-based measurement algorithm to estimate total sleep time. Many apps can also record sleep patterns by monitoring movement using a smartphone placed on the bed and, similar to wearables, monitor sleep through a microphone and acoustics. Other devices can be placed under the mattress or at the bedside to monitor pressure, movement, or sound.

Overall Aim

Several studies have shown promising evidence that support the feasibility, acceptability, and limited effectiveness of digital interventions for behavior change [13-19]. However, many of these studies have been conducted for short durations and on small sample sizes [13]. Through the use of wearable devices, this initiative aims to increase students' awareness about the importance of activity and good sleep habits through self-tracking and to develop and provide interventions and tools to help achieve healthy behaviors during and following college. We are currently conducting the first of 2 pilot studies to assess student interest, adherence, feasibility, and staffing needs and to gather baseline data through wearables and surveys on stress, diet, physical activity, and sleep behaviors. The pilot data will inform the larger launch of interventions tailored per student behaviors for the entire freshman class.

Methods

Design

Although there are many types of behaviors to target, we initially focused on activity and sleep, as poor sleep habits have a profound waterfall effect on not only health but also on academic and social elements [20-22]. Poor quantity and quality of sleep can affect individuals of all ages, gender, and racial and ethnic backgrounds. Many adults fall short of the recommended goal of 7 or more hours of daily sleep [23], with approximately 35% of US adults reporting insufficient sleep [24,25]. In college students, higher rates of insufficient sleep have been reported, with one study reporting 70% of students with inadequate sleep [26,27].

To our knowledge, this type of an initiative has not been implemented in the United States in a campus-wide setting. Preparing for such a large endeavor involved discussions and collaborations with 3 major groups on campus: (1) university leadership and administrators, (2) information technology staff, and (3) students. Given the focus on students, support from the university administration was critical to the development and

implementation of such a large initiative. We convened discussions with campus administrators in student health, student wellness, and student affairs to identify support and begin to outline the initiative and develop a proposal. With their commitment, we applied and received funding from the Office of the Provost.

We proposed a 3-year plan to develop and launch this initiative to an entire freshman class, which has been described in detail in the following sections. Specifically, the 3-year proposal included a year-long planning process, followed by 2 year-long pilot studies to assess the feasibility and acceptability of wearables and the impact of connecting students to campus interventions to promote healthy behaviors.

Year 1: Planning Phase

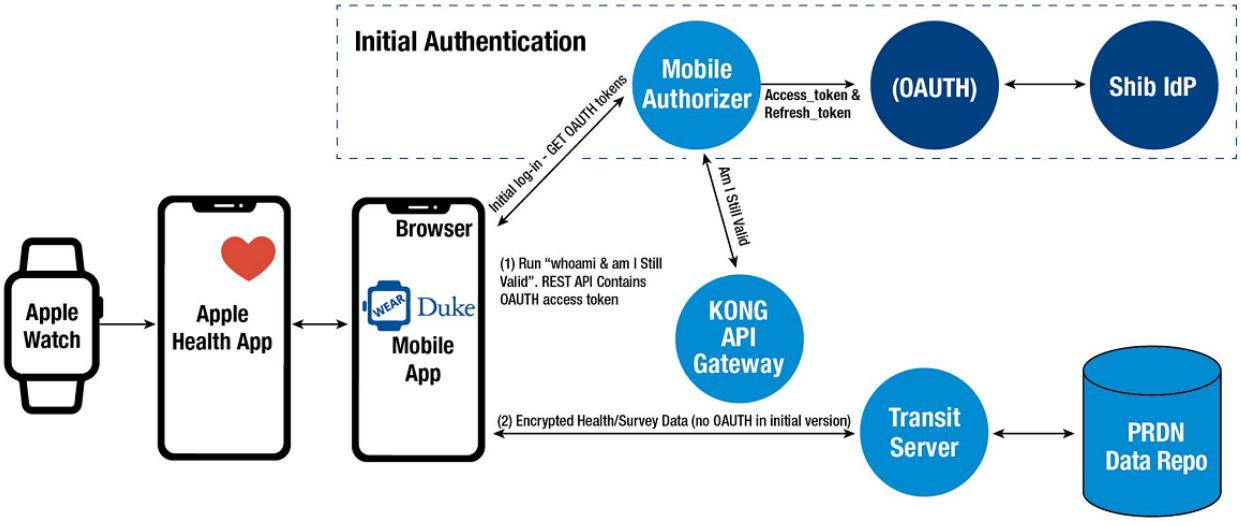
During the planning phase, we assembled an interdisciplinary faculty advisory team with expertise in mHealth technologies, app development, health behaviors, data science, psychology, sleep disorders, and student affairs and student health with continued participation from campus administrators. The faculty team advised the principal investigators on the development, infrastructure, and implementation of the initiative and is being briefed monthly. To successfully develop and implement an initiative intended for students, we considered it essential to involve students at every stage of planning, development, and implementation. Thus, in parallel with establishing a faculty advisory team, we established a student advisory committee. The student team was responsible for gathering feedback from students about their interest and concerns about the initiative; leading the selection of the name, logo design, and wearable; informing the development of an iOS app; creating the website; and outlining the protocol for the initiative (the student team was supported by the Duke Bass Connections program).

Information Technology Solution

With the proposed use of digital technologies and the collection of multiple types of data, we engaged with the Duke's Office of Information Technology (OIT) staff to develop a strategy for managing the full lifecycle of app development, deployment, and data collection processes. Our team worked very closely with software developers from the campus to develop the infrastructure for data collection, analysis, secure data storage, and technical support for the students. This included the development of a study app and website as well. Furthermore, the OIT staff are instrumental in working with the institutional review board and security office to ensure that the study meets all security and regulatory requirements.

To receive data from Apple Watch (Apple Inc) and push weekly surveys to participants, we developed a study app with OIT. For the pilot studies, we only developed an iPhone Operating System (iOS) app given that a large proportion of incoming students were estimated to have Apple iPhones based on the current student body data (>90%). This included creating a secure server architecture that allowed us to retrieve and store data for analyses (Figure 1). Access to the WearDuke app through the University App Store was restricted to students who finished the consent process. When installing the WearDuke app, students were asked to allow the app to access each category of study data stored in Apple HealthKit (eg, heart rate, steps). HealthKit is an in-built app that aggregates health-related data on the iPhone. Data from Apple Watch is transferred and stored on HealthKit. The WearDuke app retrieves and encrypts the permitted data from HealthKit and then transfers the encrypted data to a secure server on Duke's protected network. Any third-party app or any other device connected to the phone that is providing data to the HealthKit will be available to the WearDuke app and collected and stored with the study data. The source of the data (Apple Watch, app, or other device) is included with the data collected. Survey data are also collected in the WearDuke app and encrypted and transferred in the same manner as the data from Apple HealthKit.

Figure 1. WearDuke iOS architecture. OAUTH: open authorization; Shib IdP: Shibboleth identity provider; KONG: Kong API Gateway; API: Application programming interface; REST API: Representational state transfer API; PRDN: Protected research data network; Repo: Repository; App: Application.



Student Focus Groups

An undergraduate student team conducted a series of focus groups with undergraduate students to ascertain the general interest about the initiative, the name and logo, the familiarity with wearables, the features of a companion app, the likelihood to complete surveys and wear wearables, and an incentive system. We will continue to have undergraduates involved in the initiative throughout the pilot year to help guide and prepare for the second pilot study.

Wearable Selection

For this initiative, we will provide the enrolled students with a choice of wearable to monitor activity and heart rate; sleep behavior data will be gathered through an app. With our student team, we reviewed available wearables based on several criteria: type (wrist or ring), battery life, measurements, style, other features of interest to students, and methods of data access. We were unable to identify wearables that are compatible with both Android and iOS smartphones, have a comprehensive set of desired measures (sleep and activity) that are clinically accurate, and possess other features that are attractive to a college student population. Thus, we decided to offer the Fitbit Charge 3 and Apple Watch 3 to Android and iOS smartphone users for the first pilot study, respectively. We will re-evaluate the choice of wearables for the second pilot, informed by the first pilot study and newly introduced wearable models.

Year 2: Pilot Study 1

In the pilot feasibility study, we worked with the campus institutional review board, information security and privacy offices, and information technology staff to create an informed consent process that would be both understandable and transparent to young adults. We worked with Duke Web Services in OIT to create a university-approved website for the study.

Recruitment and Enrollment

In July 2019, we invited students from a single residence hall (N=175) to participate in the study to assess the feasibility and acceptance of wearables to measure health behavior and to identify trends in health behaviors over the students' freshman year. To be eligible for the study, students must be enrolled as a full-time freshman (Class of 2023), residing in the selected residence hall approved for the study; college freshmen residing in other dorms were not eligible to participate. Students received a letter in the post and by email about the study, inviting them to enroll after reviewing the website; adult-aged students interested in enrolling completed an electronic informed consent form. Students must acknowledge reviewing each section of the electronic consent form (in the checkbox at the end of each section) and sign the form. For students aged under 18 years, parental consent was obtained first, followed by student informed assent (identical in content to the student consent). We worked with OIT to develop a process using institutional data to ensure that only students who met the eligibility criteria were able to be included in the study—a solution using group management that includes information about students who turn 18 after enrolling (with parental consent) to ensure that we are collecting data from properly consented enrollees. Participants were required to attend an in-person orientation the week before the fall semester began; study investigators described the study, benefits, and risks before distributing the wearables. Assistance was provided as needed to set up the wearable and connect to their phones and install and set up the app. OIT support was available throughout the study to address student issues with the wearable or companion app. A WearDuke email was established to facilitate communication with the students.

Pilot Study 1: Methodology and Measures

For the first pilot study, we are primarily interested in ascertaining feasibility, student interest, and student experiences as well as infrastructure needs. Feasibility will be measured with study records and feedback regarding the number enrolled, number of withdrawals, frequency of wearing the wearable daily, survey completion rate, and use of information technology support services. In addition to the data on wearables (Table 1), we will administer weekly surveys to gather more information about students' sleep habits, caffeine use, overall health and mental health, and academic performance (Table 2). We administered a baseline/demographic survey following enrollment. Students with an iOS smartphone were asked to install the developed companion app to complete surveys. Students with Android-based smartphones use the Fitbit app and will complete surveys through Research Electronic Data Capture (REDCap), an electronic data capture tool hosted at Duke University [28,29]. REDCap is a secure, web-based software platform designed to support data capture for research studies. All pilot data will be stored on internal secure servers within the university. We will assign each participating student a unique study identification number. We will review and apply what we learned from year 1 and begin to develop intervention components for year 2.

To remain enrolled in the study, students must maintain an active participation status in the study through the completion of half of the surveys (2 per month) and wearing the watch a minimum of 3 days weekly. Adherence is determined through the analysis of heart rate data from the raw wearable data; an hour worn is defined by collecting at least one heart rate sample from the wearable within that hour. Thus, the *set* of hours of all heart rate samples within a day is defined to be the number of hours the user has worn the watch for that day (ie, the hours in which the samples occurred are 0, 2, 2, 13, 13, 14, 15, which is 5 hours worn). The student must wear the watch for 8 hours to be counted as worn. Incentives were provided to students who completed weekly surveys and wore the wearable for at least the minimum number of days weekly.

Table 1. Mobile device data collection.

Data type	Fitbit Charge 3	Apple Watch 3
Activity/exercise	Daily: steps taken, distance, and floors climbed; minutes lightly, fairly, and very active; minutes sedentary	Daily: steps, distance, flights climbed, exercise time, and stand hours
Energy	Daily: calories burned	Daily: basal and active energy burned
Heart rate	Daily average: heart rate and heart rate zone	Daily average: resting heart rate, walking heart rate, and heart rate variability
Sleep analysis	Daily: time asleep and sleep stages (Rapid eye movement, light, and deep)	Native Apple app (clock) or other third-party app (student's choice)

Table 2. Weekly surveys.

Subject/topic	Survey instrument
Demographics	• Race, gender, and campus activities
Sleep (habits, environment, and circadian preference)	• Sleep/wake behavior problems • Use of sleep aides (medications, music, and blinders; adapted from National Sleep Foundation [30]) • Sleep quality (Patient-Reported Outcomes Measurement Survey (PROMIS) Sleep Practices and Sleep Disturbance) and daytime sleepiness (PROMIS SRI) [31,32] • Circadian preference (Morningness-Eveningness Reduced Questionnaire [33]) • Physical environment/number of roommates
Mental health	• Depression (Center for Epidemiologic Studies Depression scale) [34] • Stress (Cohen's Perceived Stress Scale [35])
Nutrition (habits and knowledge)	• Nutrition knowledge and habits (adapted from [36]) • Caffeine intake [37]
General health status	• General physical/mental health status [38] • Reported number of absences owing to sickness • Reported number of visits to student health
Academic schedule/performance	• Intended major/minor • Courses/schedule • Academic performance (GPA) • Average hours per week for campus activities, work study, and/or employment
Physical activity	• Physical activity habits (adapted from [36]) • Participation in campus programs
Experience with wearables and apps	• Use of sleep tools/alerts/tracking • Experience with wearables and health-related apps

Year 3: Pilot Study 2

Data collected from the first pilot will inform an expanded, second pilot study that will focus on the evaluation of interventions to promote healthy behaviors. In this second larger study, we will implement an improved app and incentive strategy, continue to monitor activity and sleep behavior, and evaluate the uptake and effectiveness and use of campus activities and interventions. Our intervention will be guided by theory-driven behavior change principles that leverage capabilities of continuous monitoring technologies and address the unique preference of individuals. For this study, we will use the Healthy Apps 4 M's conceptual model of monitoring, modeling, motivating, and modifying [39]. These principles will help guide our intervention development that will allow for near real-time interventions based on continually observed behavioral response data.

Educational interventions may include workshops, in-person or web-based guest speakers, and sending out healthy tips via push notifications. Physical activity interventions will be available for students at all levels and experiences, individual and group based. Sleep interventions may include both physical sleep aides (eg, pillows) or guidance for establishing a nightly routine to enable adequate sleep quality and quantity. For example, students may benefit from learning how to de-stress, perhaps through simply reducing mobile phone use before going to sleep (it is estimated that about 40% of Americans, including 72% of adolescents, use mobile technology before going to sleep [40,41]). Ideally, we will alert students to various interventions based on their preferences and wearable data. Assuming successful completion of the pilot studies, we hope to expand the initiative to the entire incoming class.

As with the first pilot study, we will re-assess student interest and experiences through the enrollment rate, drop-out rate, daily wearable adherence, damage to wearables, survey completion rate, and use of information technology support services. In addition to objective data collected from the wearables, we will administer weekly surveys to gather more information about students' sleep habits, caffeine use, overall health and mental health, and academic performance. The second pilot will introduce app notifications based on student preferences and wearable data for activities, workshops, or other resources to promote well-being.

Data Analysis

This is an exploratory study, and there are no interventions or hypotheses. As these pilot studies are intended to primarily assess feasibility and acceptance, the project is not powered to detect statistically significant effects. Descriptive statistics and some nonparametric tests, eg, Wilcoxon signed rank tests, Fisher exact tests, and other appropriate measures will be calculated using the R environment. We will assess student variables such as gender and intended major with regard to the survey and wearables data. Depending on the scoring algorithm for each survey instrument, any missing data may prohibit the generation of a score. In other cases, descriptive statistics are generated per question, and "prefer not to answer" is a part of the range of responses quantified. Working with Duke OIT, we can provide our analytics staff with tools such as R-Studio, as well as state of the art Python-based tools for the analysis of very large datasets.

Educational Opportunities

In addition to promoting healthy lifestyles for students and a healthy campus culture, we envision that students and faculty will have the opportunity to conduct research with anonymized datasets, in turn providing students with a real-world dataset to gain skills in research and data analysis. We hope to have the opportunity to work with students to develop smaller studies within the larger group of WearDuke participants to address or gather data for other areas not currently addressed. Through this experience, we anticipate that they will learn about study design, generating hypotheses, survey development and other research methodology, human subject protections, and the analysis of complex datasets. Students may also work on improvements to the app or the development of new campus interventions to promote healthy behaviors.

The study has also allowed us to work with classes across campus such as computer science and engineering. For example, an Android development class created a prototype for the WearDuke study. Not only did this allow students to learn how to work with clients and develop a product for a real-world project but also it provided a beta version and wireframe of an Android version of the study app. We would then be able to take this student work to professional developers as a prototype that can be built upon. We expect that this educational approach will be repeated, and future classes will build an additional app or technology-related features for the study. This may include testing new digital health technologies that could measure sleep, among others.

Results

We completed the year-long planning phase and obtained approval from the Duke University Campus Institutional Review Board in June 2019. The website was made publicly accessible in July 2019 [42]. The first pilot study was launched in July 2019 in a single freshman residence hall. An electronic consent process was established, and all students attended a mandatory orientation session on August 24 to 25, 2019. A total of 175 students were eligible to participate, and 120 (68%) students consented to participate, and 114 (65%) students started the study (see Figure 2). Student demographic data are presented in Table 3.

All students were required to complete a demographic survey during the enrollment process. A survey was administered each week of the semester (16 total surveys). Weekly survey completion rates ranged from a high of 95% to 70% (Figure 3). By the end of the fall semester, a total of 8 students withdrew from the study or were withdrawn by study staff owing to inactive participation. Watch replacements were provided for a total of 10 students for watches that were lost (4), broken (5), or nonfunctional (1). An analysis for each survey and wearables dataset has commenced. In addition to generating summary statistics, we will test for differences between major student demographic features such as gender and intended major for both survey and wearables data as well as test for changes in behaviors across the semester in repeated measures. We aim to publish the initial results of the first pilot study in fall 2020.

Figure 2. CONSORT Flowchart of WearDuke initiative through end of fall semester.

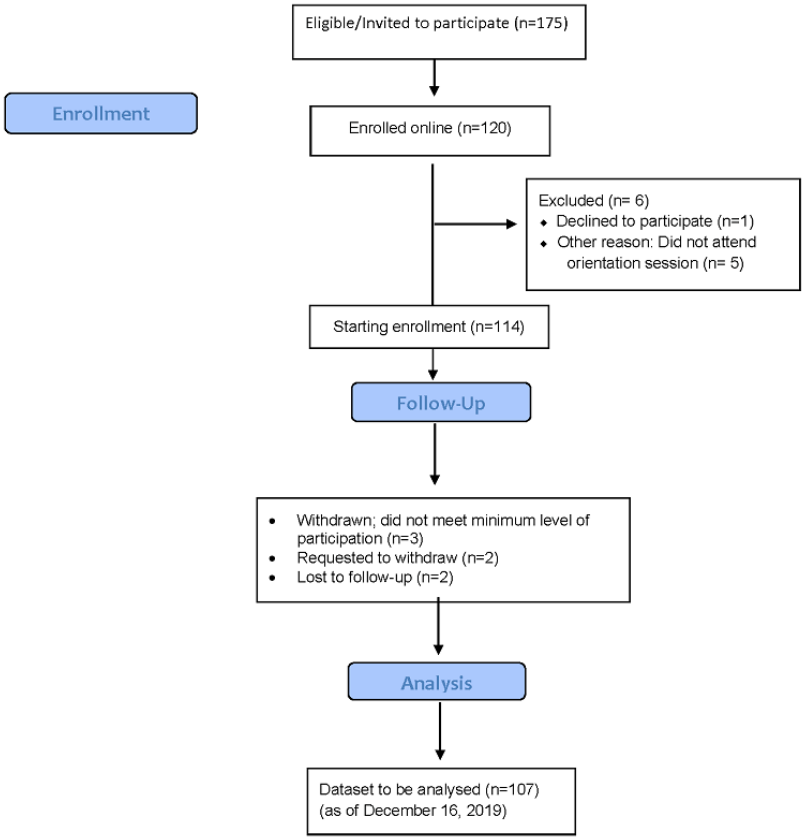
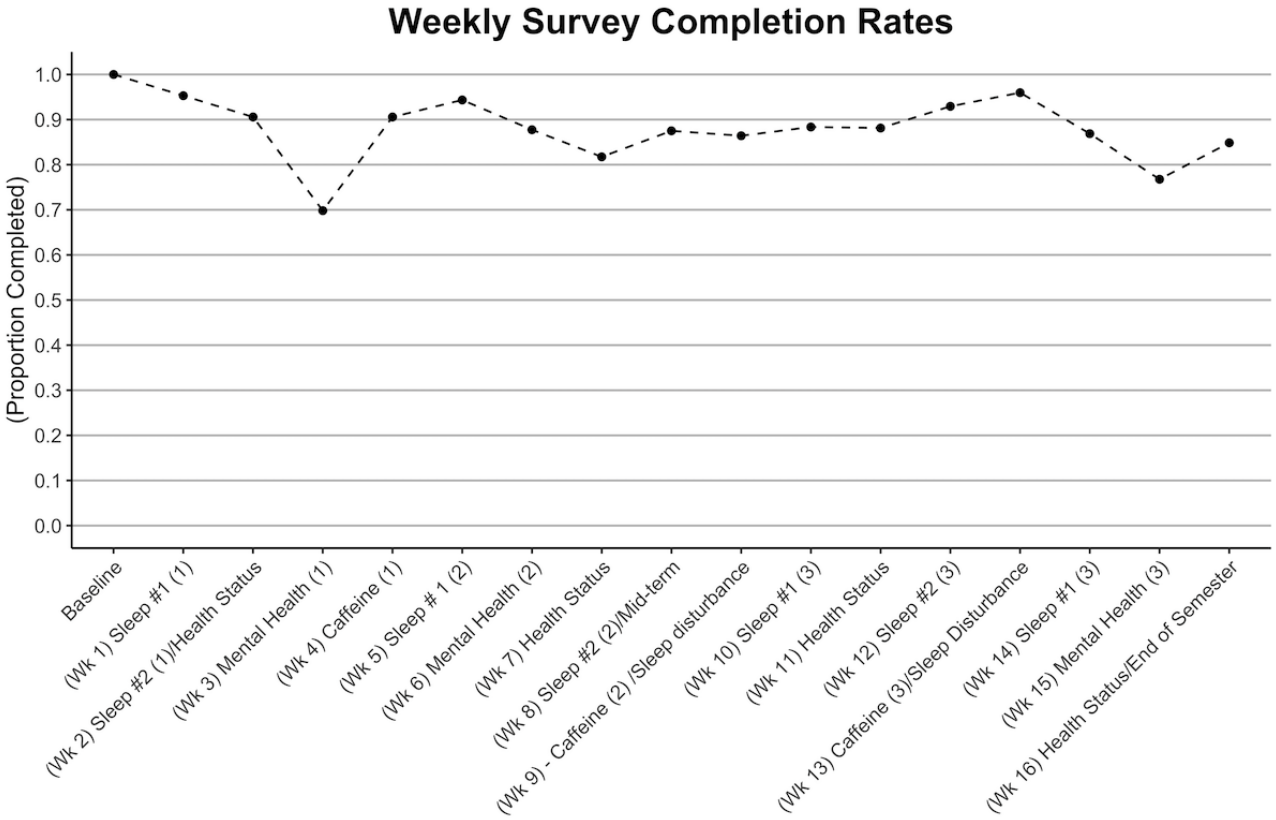


Table 3. Participant demographic data (N=114).

Characteristics	Values, n (%)
Male	62 (54.3)
Age (years)	
<18	13 (11.4)
18	95 (83.3)
19	5 (4.4)
20 or older	1 (0.88)
Hispanic	15 (13.1)
Race	
White	63 (55.3)
African American	8 (7.0)
Asian	27 (23.7)
Other	2 (1.8)
More than 1 race	13 (11.4)
School	
Trinity School of Arts & Sciences	85 (74.6)
Pratt School of Engineering	29 (25.4)
Varsity sports	2 (1.8)
Own an iPhone	106 (93.0)

Figure 3. Weekly survey completion rate for Fall 2019 semester for Apple Watch participants.



Discussion

Preliminary Findings

In a time of rapid growth of digital technologies in the educational and health space, we seek to promote familiarity with these technologies to improve health and well-being for undergraduate students. No other informal learning opportunity exists for students on campus to gain experience with these technologies, and the potential for new educational opportunities further expands the multiple benefits that this initiative may yield. In the first of the 2 pilot studies commenced, the study launched with 65% of eligible students. Halfway through the study (at the end of the fall semester), the attrition rate was 7%, and the average weekly survey completion rate was 88%.

The transition to college is a period when new and long-term habits are being formed. With ongoing concerns about the health and well-being of college students today, we are developing a tool to help students adjust to the collegiate environment and emphasize the need for students to make time for themselves and establish healthy lifestyles. Such an expansive undertaking requires engagement with many stakeholders and, most importantly, with the targeted student population. We have begun discussions to expand the initiative to the health system and other groups on campus including graduate students, residents/trainees, staff, and faculty. We hope our initial experiences will inform the use of digital technologies in other settings.

We do acknowledge several limitations to this study. The first is the limited sample size of the first pilot, which did not allow us to draw statistical conclusions from the data. The first pilot also does not test the effect of any intervention, so we cannot assess the impact on health and wellness outcomes of using these technologies. Students may become tired or lose enthusiasm to complete the weekly surveys, particularly the repeated surveys, limiting our ability to detect trends over and between semesters. Finally, this study is at a single college campus and does not reflect the diversity of college campuses and student bodies across the United States. We also recognize the economic limitations of broadly implementing an mHealth-based initiative and its general feasibility in other educational settings (or settings with shared living arrangements) or for population health initiatives [43,44].

Conclusions

College students' health and well-being has garnered widespread attention and concern in recent years. Similarly, the expansion and evaluation of digital technologies has grown in recent years for different target populations. We believe such an initiative will have both short-term and long-term implications for students in learning about and facilitating healthy behaviors that will optimize well-being and academic performance throughout college as well as establish healthy behaviors that will have a lasting impact on their health and well-being after college. Furthermore, the data and experiences from this initiative can inform the development of similar programs in other educational and group-based settings (eg, military and nursing homes) to improve residents' overall health in settings that are new/unfamiliar, stressful, and/or resource-limited.

Acknowledgments

This study was funded by a grant from the Duke Office of the Provost, with additional support from the Duke School of Medicine, Duke Health System, Duke Mobile App Gateway via the Duke Clinical and Translational Science Institute (US National Institutes of Health grant UL1TR002553), and private donors. The student team is supported by the Duke Bass Connections program. The authors would like to thank the students and staff who have generously contributed their time toward the development of this initiative.

Conflicts of Interest

None declared.

References

1. Calhoun BH, Maggs JL, Loken E. Change in college students' perceived parental permissibility of alcohol use and its relation to college drinking. Addict Behav 2018 Jan;76:275-280 [FREE Full text] [doi: 10.1016/j.addbeh.2017.08.025] [Medline: 28886576]
2. Poulos NS, Pasch KE. Energy drink consumption is associated with unhealthy dietary behaviours among college youth. Perspect Public Health 2015 Nov;135(6):316-321. [doi: 10.1177/1757913914565388] [Medline: 25667166]
3. Liu CH, Stevens C, Wong SH, Yasui M, Chen JA. The prevalence and predictors of mental health diagnoses and suicide among US college students: implications for addressing disparities in service use. Depress Anxiety 2019 Jan;36(1):8-17 [FREE Full text] [doi: 10.1002/da.22830] [Medline: 30188598]
4. Lee BL, Jordan HR, Madson MB. The moderating effects of college stress on the relationship between protective behavioral strategies and alcohol outcomes. Subst Use Misuse 2019;54(11):1845-1852. [doi: 10.1080/10826084.2019.1618330] [Medline: 31240985]
5. Vilaro M, Colby S, Riggsbee K, Zhou W, Byrd-Bredbenner C, Olfert M, et al. Food choice priorities change over time and predict dietary intake at the end of the first year of college among students in the US. Nutrients 2018 Sep 13;10(9):pii: E1296 [FREE Full text] [doi: 10.3390/nu10091296] [Medline: 30217004]
6. Pew Research Center. Washington, D.C: Pew Research Center; 2019 Jun 12. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-09-30]
7. DeMasi O, Feygin S, Dembo A, Aguilera A, Recht B. Well-being tracking via smartphone-measured activity and sleep: cohort study. JMIR Mhealth Uhealth 2017 Oct 5;5(10):e137 [FREE Full text] [doi: 10.2196/mhealth.7820] [Medline: 28982643]
8. Þórarinsdóttir H, Kessing LV, Faurholt-Jepsen M. Smartphone-based self-assessment of stress in healthy adult individuals: a systematic review. J Med Internet Res 2017 Feb 13;19(2):e41 [FREE Full text] [doi: 10.2196/jmir.6397] [Medline: 28193600]
9. Swarthout M, Bishop M. Population health management: review of concepts and definitions. Am J Health Syst Pharm 2017 Sep 15;74(18):1405-1411. [doi: 10.2146/ajhp170025] [Medline: 28887342]
10. Shaw RJ, Bonnet JP, Modarai F, George A, Shahsahebi M. Mobile health technology for personalized primary care medicine. Am J Med 2015 Jun;128(6):555-557. [doi: 10.1016/j.amjmed.2015.01.005] [Medline: 25613298]
11. Dunn EE, Gainforth HL, Robertson-Wilson JE. Behavior change techniques in mobile applications for sedentary behavior. Digit Health 2018;4:2055207618785798 [FREE Full text] [doi: 10.1177/2055207618785798] [Medline: 31463076]
12. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
13. Badawy SM, Kuhns LM. Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: a systematic review. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]
14. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered eHealth interventions for children, adolescents, and adults with sickle cell disease: systematic review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
15. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
16. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
17. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]

18. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
19. Badawy SM, Thompson AA, Kuhns LM. Medication adherence and technology-based interventions for adolescents with chronic health conditions: a few key considerations. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e202 [FREE Full text] [doi: 10.2196/mhealth.8310] [Medline: 29273573]
20. Amaral AP, Soares MJ, Pinto AM, Pereira AT, Madeira N, Bos SC, et al. Sleep difficulties in college students: The role of stress, affect and cognitive processes. Psychiatry Res 2018 Feb;260:331-337. [doi: 10.1016/j.psychres.2017.11.072] [Medline: 29227897]
21. Becker SP, Dvorsky MR, Holdaway AS, Luebbe AM. Sleep problems and suicidal behaviors in college students. J Psychiatr Res 2018 Apr;99:122-128 [FREE Full text] [doi: 10.1016/j.jpsychires.2018.01.009] [Medline: 29448236]
22. Gawlik K, Melnyk BM, Tan A, Amaya M. Heart checks in college-aged students link poor sleep to cardiovascular risk. J Am Coll Health 2019;67(2):113-122. [doi: 10.1080/07448481.2018.1462823] [Medline: 29652617]
23. Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, Buysse D, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the american academy of sleep medicine and sleep research society. Sleep 2015 Jun 1;38(6):843-844 [FREE Full text] [doi: 10.5665/sleep.4716] [Medline: 26039963]
24. Centers for Disease Control and Prevention. Atlanta, GA Sleep and Sleep Disorders URL: https://www.cdc.gov/sleep/index.html [accessed 2019-09-30]
25. National Center for Health Statistics. Centers for Disease Control and Prevention. Atlanta, GA Table SLP–2a. Age-adjusted percentage (with standard errors) of adults aged 18 and over who met the Healthy People 2020 objective for sufficient sleep, by sex and selected characteristics: United States, annualized, 2011–2014 URL: https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2011-2014_AHB_Table_SLP-2.pdf [accessed 2019-09-30]
26. Hershner SD, Chervin RD. Causes and consequences of sleepiness among college students. Nat Sci Sleep 2014;6:73-84 [FREE Full text] [doi: 10.2147/NSS.S62907] [Medline: 25018659]
27. Orzech KM, Salafsky DB, Hamilton LA. The state of sleep among college students at a large public university. J Am Coll Health 2011;59(7):612-619. [doi: 10.1080/07448481.2010.520051] [Medline: 21823956]
28. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
29. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
30. National Sleep Foundation. 2004 Sep. 2005 Sleep in America Poll URL: https://www.sleepfoundation.org/sites/default/files/inline-files/SIAQuestionnaire2005.pdf [accessed 2019-09-30]
31. Health Measures. PROMIS - Obtain and Administer Measures URL: http://www.healthmeasures.net/explore-measurement-systems/promis/obtain-administer-measures [accessed 2019-09-30]
32. Hanish AE, Lin-Dyken DC, Han JC. PROMIS sleep disturbance and sleep-related impairment in adolescents: examining psychometrics using self-report and actigraphy. Nurs Res 2017;66(3):246-251 [FREE Full text] [doi: 10.1097/NNR.0000000000000217] [Medline: 28448375]
33. Adan A, Almirall H. Horne & Östberg morningness-eveningness questionnaire: A reduced scale. Pers Individ Dif 1991;12(3):241-253 [FREE Full text] [doi: 10.1016/0191-8869(91)90110-W]
34. Radloff LS. The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. J Youth Adolesc 1991 Apr;20(2):149-166. [doi: 10.1007/BF01537606] [Medline: 24265004]
35. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
36. National Cancer Institute. Division of Cancer Control and Population Sciences. Bethesda, MD Family Life, Activity, Sun, Health, and Eating (FLASHE) study URL: https://cancercontrol.cancer.gov/brp/hbrb/flashe.html [accessed 2019-09-30]
37. Landrum RE. College students' use of caffeine and its relationship to personality. Coll Stud J 1992;26(2):151-155 [FREE Full text]
38. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System: BRFSS Questionnaires URL: https://www.cdc.gov/brfss/questionnaires/index.htm [accessed 2019-09-30]
39. Spring B, Gotsis M, Paiva A, Spruijt-Metz D. Healthy apps: mobile devices for continuous monitoring and intervention. IEEE Pulse 2013;4(6):34-40 [FREE Full text] [doi: 10.1109/MPUL.2013.2279620] [Medline: 24233190]
40. Orzech KM, Grandner MA, Roane BM, Carskadon MA. Digital media use in the 2 h before bedtime is associated with sleep variables in university students. Comput Human Behav 2016 Feb;55(A):43-50 [FREE Full text] [doi: 10.1016/j.chb.2015.08.049] [Medline: 28163362]
41. National Sleep Foundation. Annual Sleep in America Poll Exploring Connections with Communications Technology Use and Sleep URL: https://sleepfoundation.org/media-center/press-release/annual-sleep-america-poll-exploring-connections-communications-technology-use-(doi: [accessed 2019-09-30]

42. WearDuke. URL: https://wearduke.duke.edu/ [accessed 2020-02-12]
43. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
44. Badawy SM, Kuhns LM. Economic evaluation of text-messaging and smartphone-based interventions to improve medication adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]

Abbreviations

mHealth: mobile health
OIT: Office of Information Technology
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 02.10.19; peer-reviewed by L Piwek, S Badawy; comments to author 18.11.19; revised version received 21.12.19; accepted 27.01.20; published 03.04.20.

Please cite as:
Haga SB, Shaw R, Kneifel C, Bond SJ, Ginsburg GS
Promoting Wellness Through Mobile Health Technology in a College Student Population: Protocol Development and Pilot Study
JMIR Res Protoc 2020;9(4):e16474
URL: https://www.researchprotocols.org/2020/4/e16474
doi:10.2196/16474
PMID:32242827

Protocol

Reducing Drinking Among People Experiencing Homelessness: Protocol for the Development and Testing of a Just-in-Time Adaptive Intervention

Michael S Businelle[1*], PhD; Scott T Walters[2*], PhD; Eun-Young Mun[2], PhD; Thomas R Kirchner[3], PhD; Emily T Hébert[1], DrPH; Xiaoyin Li[2], PhD

[1]Oklahoma Tobacco Research Center, Stephenson Cancer Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States
[2]School of Public Health, University of North Texas Health Sciences Center, Fort Worth, TX, United States
[3]School of Global Public Health, New York University, New York City, NY, United States
[*]these authors contributed equally

Corresponding Author:
Michael S Businelle, PhD
Oklahoma Tobacco Research Center
Stephenson Cancer Center
University of Oklahoma Health Sciences Center
655 Research Parkway
Oklahoma City, OK, 73104
United States
Phone: 1 4052718001 ext 50460
Email: michael-businelle@ouhsc.edu

Abstract

Background: Adults who are homeless are more likely to have alcohol use disorders (AUDs) compared with domiciled adults. Although AUD treatments are commonly available, many factors (eg, transportation limitations and inability to schedule appointments) contribute to low treatment completion rates and low success rates of these interventions among adults experiencing homelessness. Most adults who are homeless own mobile phones; however, no interventions have been developed that use mobile devices to deliver and support AUD interventions for this population. Mobile phone–based AUD interventions may reduce barriers that have limited the use and utility of traditional interventions.

Objective: The aim of this study is to (1) identify variables (eg, affect, stress, geolocation, and cravings) that predict drinking among homeless adults (phase I), (2) develop a mobile intervention that utilizes an algorithm to identify moments of risk for drinking and deliver treatment messages that are tailored to the individual's current needs in real time (phase II), and (3) pilot test the intervention app (phase III).

Methods: In phase I, adults experiencing homelessness with an AUD (N=80) will complete baseline, equipment, 2-week, and 4-week follow-up visits in person. Participants will be prompted to complete five daily ecological momentary assessments on a study-provided smartphone for 28 days. The smartphone app will collect GPS coordinates every 5 min for the entire 28-day study period. Participants will wear a transdermal alcohol sensor that will objectively measure alcohol use. In phase II, we will use phase I data to develop an algorithm that identifies moments of heightened risk for drinking and develop treatment messages that address risk factors for drinking. Phase III will pilot test the intervention in 40 adults experiencing homelessness with AUD.

Results: This project was funded in June 2018. IRB approval was obtained in October 2018, and data collection for phase I began in February 2019. Phase III data collection is expected to conclude in 2020. To date, 80 participants have consented to the study, and data analysis for phase I will begin in early 2020.

Conclusions: This research will highlight intervention targets and develop a novel intervention for understudied and underserved adults experiencing homelessness with AUD.

International Registered Report Identifier (IRRID): DERR1-10.2196/15610

(JMIR Res Protoc 2020;9(4):e15610) doi:10.2196/15610

KEYWORDS

alcohol use disorder; mobile health; smartphone; ecological momentary assessment; homeless persons

Introduction

Background

An estimated 6.2% of US adults will be homeless at some point in their lifetime [1]. Adults experiencing homelessness have higher rates of disease, greater risk of interpersonal violence, shorter life expectancies, and disproportionately higher health care utilization and costs compared with housed individuals [2-5]. Adults experiencing homelessness are 8 times more likely than adults in the general population to be alcohol dependent [6,7], and the high prevalence (29%-63%) [8-11] of alcohol use disorders (AUDs) is a leading contributor to the higher morbidity and mortality among adults experiencing homelessness. For example, one study found that alcohol was responsible for an estimated 17% of all deaths among homeless adults in Boston, a rate 6 to 10 times greater than in the general population [12].

Although shelter-based AUD treatments are common [13], adherence is typically poor [6]. Homeless individuals experience a number of barriers to receiving AUD treatment, including lack of stable housing [14], fractured social networks [15], and high rates of co-occurring problems [16]. There is evidence that shelter-based interventions are feasible [17] and can decrease drinking [18]; however, keeping clients engaged in treatment is a challenge [19,20]. In an analysis of 15 substance use disorder (SUD) treatment programs for homeless individuals funded by a National Institute on Alcohol Abuse and Alcoholism cooperative agreement [21], dropout rates ranged from 67% to 97.5%. Not a single program was completed by more than one-third of clients. Major reasons for dropout included poor client motivation, a desire to reconnect with family or friends outside of treatment, dissatisfaction with the program structure or environment, and other logistical difficulties.

Relatively little is known about the environmental, cognitive, affective, and behavioral antecedents of alcohol use in adults experiencing homelessness. Alcohol use has most often been examined using traditional lab-/clinic-based assessment methods that are not well suited to capturing the complicated street-level interactions experienced by most adults experiencing homelessness [22,23]. Ecological momentary assessment (EMA), in which handheld devices (eg, smartphones) are used to capture moment-to-moment experience via brief surveys, reduces recall bias and more accurately measures antecedents and correlates of alcohol use in natural settings [22,24-26]. In addition, recent technological advances in smartphone sensors have made it possible to passively collect continuous geolocation data (ie, GPS coordinates) alongside EMA [27]. Thus, momentary changes in key variables can be tracked, studied, and potentially used to initiate real-time interventions and engage clients in treatment.

Multiple studies have identified momentary predictors of smoking relapse [28-33] and have shown that the composite scores based on daily diaries of self-efficacy and motivation are more consistent predictors of drinking outcomes than global measures of self-efficacy and motivation among problem drinkers [34]. Thus, EMAs may be better suited than traditional clinic-based, trait-like measures to identify mechanisms that drive alcohol use.

Although EMA has been used in a variety of populations and for multiple health outcomes, only a few studies [33,35] have collected EMA data in adults experiencing homelessness. Furthermore, few studies have used both EMA and geolocation to assess risk for alcohol use despite the fact that research studies have indicated that most adults experiencing homelessness possess phones with active service (ie, we have conducted two large survey studies of adults experiencing homelessness (Dallas and Oklahoma City) and found that 58.4% of 394 surveyed adults in Dallas (2013) and 71.9% of 589 surveyed adults in Oklahoma City (2016) possessed cellular phones with active service). We are not aware of any studies that have combined this information to estimate risk and intervene in real time to reduce drinking among homeless adults. We believe combining EMA and geolocation data will help improve our understanding of the mechanisms that lead people to drink and pave the way toward more effective and cost-effective alcohol treatments for this high-risk group. This paper describes the rationale and design for a three-phase treatment development study to develop a just-in-time adaptive smartphone intervention (JITAI) to reduce alcohol use in adults experiencing homelessness.

Objectives

During phase I, we will use smartphones and passive sensing to continuously monitor geolocation and to measure psychosocial variables (eg, negative affect, stress, and urge to drink) and alcohol use in a sample of 80 adults experiencing homelessness enrolled in shelter-based treatment programs. EMAs will be used to examine the moment-to-moment relationship between social cognitive theory (SCT) constructs (eg, affect, abstinence motivation and self-efficacy, alcohol use expectancies, and cravings) [36,37], social-ecological model constructs (eg, current proximity to previous drinking areas or alcohol outlets, social setting, and social support) [38,39], and drinking. We will also assess these constructs as trait-like variables at baseline to examine how trait and state processes interact to influence drinking behaviors. Finally, phase I participants will complete quantitative and qualitative measures at the conclusion of the study. These measures will query about things they liked and disliked about the survey app design and potential intervention components that should be included in the phase III app.

In phase II, we will use this information to develop optimized risk algorithms and develop tailored treatment messages that can be provided before anticipated alcohol use given personal, situational, and environmental triggers (eg, presence of drinking others, location, elevated positive or negative moods, and high stress).

In phase III, we will pilot test the smartphone app for utility, satisfaction, and preliminary effectiveness in another sample of 40 homeless adults enrolled in shelter-based treatments. Algorithm-driven treatment messages will be automatically

delivered at the end of EMAs. Phase III participants will complete a qualitative interview that will examine their opinions of the app design and intervention content and ways to improve the app user interface. In phases I and III, self-reported alcohol use will be validated via a transdermal alcohol sensor (ie, Secure Continuous Remote Alcohol Monitor [SCRAM], Alcohol Monitoring Systems, Inc) worn by participants.

Motivational- (eg, derived from motivational interviewing) and self-efficacy- (eg, derived from SCT) themed messages are commonly used in technology-based alcohol interventions [40]. Interventions for AUDs have often drawn from these underlying theories, but mobile interventions have the additional strength of fostering self-regulation through triggering goal salience and re-evaluation of short- versus long-term goals [41]. Recent work has indicated that smartphone apps (eg, the ACHESS app) that incorporate preloaded videos, interactive features, and weekly check-ins can reduce heavy drinking days in alcohol dependent adults [42] and college students [43]. Others have begun to use geolocation data to alert individuals with SUDs about potentially high-risk environments [44-46]. For example, some SMS text messaging interventions for AUD have focused on encouraging self-regulation and planning *before* drinking episodes [47,48]. For those who are enrolled in treatment, the messages can reinforce treatment concepts. For those who are not enrolled in treatment, messages may serve as a primary intervention (or at least a reminder of past concepts) to short-circuit alcohol use before it occurs.

Our central hypothesis is that alcohol use is strongly affected by moment-to-moment risk and protective factors, and we can use EMAs to identify and automatically intervene during moments when people are at high risk for drinking. Our hypothesis is based on preliminary findings from our own studies among homeless [33,49], justice involved [50], and socioeconomically disadvantaged safety net hospital patients [31,32]. If effective, this smartphone treatment app could significantly improve treatment engagement, drinking outcomes, and quality of life among adults experiencing homelessness with AUDs.

Methods

Setting

All phases of the project will be conducted at a large homeless shelter located in Dallas, Texas. The shelter provides multiple services, including meals, mental health and substance abuse counseling, care management, housing placement, and job readiness training to approximately 85% of all homeless adults in Dallas County each year. The shelter conducts approximately 366 new intakes each month, of which approximately 32% self-report current *problems with alcohol*. See Table 1 for characteristics of the shelter population [51].

Table 1. Characteristics of people enrolled in the Bridge Homeless Recovery Program (N=394).

Characteristics	Value
Sex (female), n (%)	111 (28.2)
Race/ethnicity, n (%)	
Black	247 (62.7)
White	102 (25.9)
Latino/Hispanic	24 (6.1)
Multiracial/other	21 (5.3)
Age (years), mean (SD)	43.9 (11.8)
Socioeconomic characteristics	
Education (years), mean (SD)	11.9 (1.8)
Insured, n (%)	93 (23.6)
Employed, n (%)	39 (9.9)

Eligibility Criteria

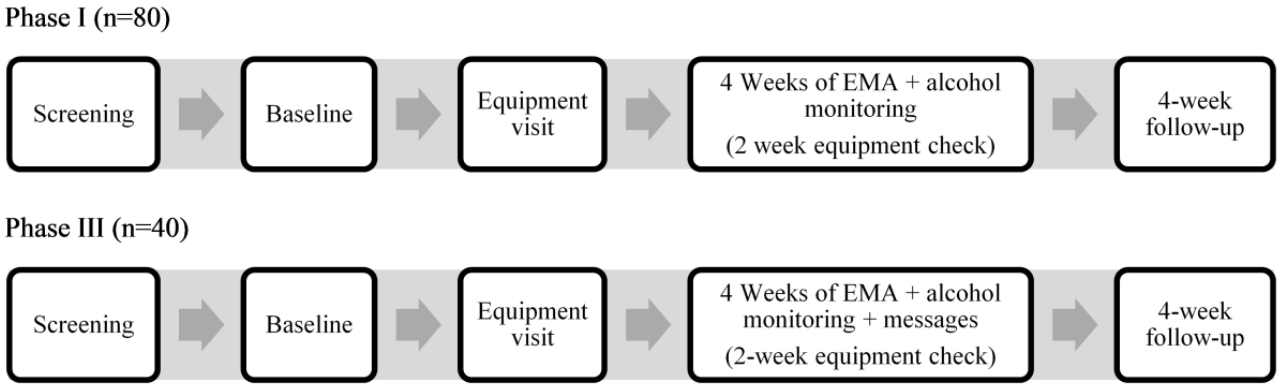
We will include relatively few exclusion criteria so that the sample will be as representative of the population as possible. Homeless individuals at the shelter will be included in the study (N=80 for phase I and N=40 for phase III) if they (1) receive a score of 8 or above on the Alcohol Use Disorders Identification Test [52] (a cutoff score suggesting hazardous and harmful alcohol use), (2) report consuming at least one standard drink of alcohol in the past week, (3) are receiving services at the shelter, (4) are willing and able to complete the baseline and follow-up visits, (5) score ≥4 on the Rapid Estimate of Adult Literacy in Medicine-Short Form indicating >6th grade English literacy level (ie, a 7th grade reading level is necessary to complete assessments), and (6) score ≥24 on the Mini-Mental State Exam indicating no substantial cognitive impairment. People with circulation problems, neuropathy, deep vein thrombosis, leg ulcers, tendonitis, diabetes, pregnancy, history of swelling, or nickel or other metal allergies will be asked to consult a shelter-based medical professional before wearing the SCRAM bracelet. Individuals will be excluded from participating if they indicate that they would be uncomfortable wearing the SCRAM bracelet for 4 weeks. Individuals will be excluded from participating in phase III if they participated in phase I.

Participant Recruitment and Procedure

Homeless adults who are receiving services at the shelter will be given a flyer that briefly outlines this study. Interested individuals will be scheduled for a screening visit to determine study eligibility. Before screening, participants will be informed that shelter services are not contingent upon study enrollment. Those who remain interested will complete the informed consent process before screening. Once eligible, participants will complete a baseline assessment, equipment visit, 2-week follow-up, and 4-week follow-up visits in a private office at the shelter. At the baseline visit, participants will complete baseline questionnaires. Approximately 3 to 7 days after the baseline assessment, participants will return for the equipment visit where they will receive the study smartphone (including instructions on how to complete the phone assessments) and be fitted with the SCRAM bracelet. All study phones loaned to participants will have a data plan which includes unlimited calls and texts and 2 GB of data per month. The app will prompt surveys, collect data, and provide intervention content even when offline (ie, no internet connection is needed for the app to work). Whenever cellular service or Wi-Fi are available, study data are automatically uploaded to the study server.

Figure 1 shows the study flow for phases I and III.

Figure 1. Phase I and III flowcharts. EMA: ecological momentary assessment.



Measures

Baseline Measures (In-Person)

In-person assessments will be administered at the baseline and 4-week visits. These measures will be used to describe the sample, identify variables that predict drinking, and help develop the treatment messages. At the baseline visit, a comprehensive locator form will be used to identify and collect multiple ways to contact participants (eg, personal email address, Facebook page, and family members' phone numbers and addresses) to reduce loss to follow-up. Data will be collected on tablet computers using Questionnaire Development System (QDS) software by NOVA Research Company (Silver Spring, Maryland). QDS utilizes a computer-administered self-interview format (ie, audio computer-assisted self-interviewing), which reduces data entry errors and the need to retain paper copies of questionnaire data. Each item appears on the computer screen while the program reads the item. Participants touch the screen to select their answers after QDS reads each item. In past studies, participants have reported few problems using the QDS program, including those with no computer experience. Trained research staff will be available to help participants who have difficulty. The baseline visit takes approximately 1 hour to complete, and the 4-week visit takes approximately 50 min to complete (see Table 2 for measures).

Table 2. In-person assessment measures.

Category	Measure
Background/history	• Locator Form • Demographic Information Questionnaire[a] • Subjective Social Status [53] • Brief Homelessness Questionnaire • Homelessness Timeline Follow-Back [54][a]
Health/mental health	• Short Form Health Survey (SF-12) [55] • Health Related Quality of Life [56] • Self-Rated Health [57] • Tobacco Questionnaire • Inadequate Sleep [58] • Time Line Follow-Back (past month alcohol) [59] • Short Inventory of Consequences [60]
Stress/affect	• Personal Victimization [61] • Perceived Stress Scale-Short Version [62] • Urban Life Stress Scale [63] • Depression [64]
Interpersonal/intrapersonal	• Interpersonal Support Evaluation List [65] • Brief Coping Orientation to Problems Experienced (COPE) [66] • Religious Participation
Treatment satisfaction	• Just-in-Time Adaptive Intervention Satisfaction Survey (quantitative and qualitative components)[b] • System Usability Scale[b] [67]

[a]Baseline only.

[b]Follow-up only.

Ecological Momentary Assessment (Phone-Based Measures)

EMA items completed on the phone (see Table 3) will assess SCT constructs (eg, affect, abstinence motivation and self-efficacy, expectancies, and cravings) and social-ecological model constructs (eg, proximity to previous drinking areas, social setting, and social support) to identify key variables and time- and location-dependent fluctuations in variables, which will be used to predict study outcomes. Most of these items have been used in our previous studies and studies from other labs [22,68]. Three types of EMAs will be used: daily diary, random sampling, and event sampling. Daily diary and random sampling EMAs will be initiated by the phone. The phone will audibly and visually cue these EMAs for 30 seconds. If the participant has not responded after 5 prompts, the assessment will be recorded as missed. Event sampling is initiated by participants if/when they consume their first drink in a day. On average, random and event sampling assessments take 2 min to complete, and daily diary assessments take less than 5 min to complete.

Daily Diary

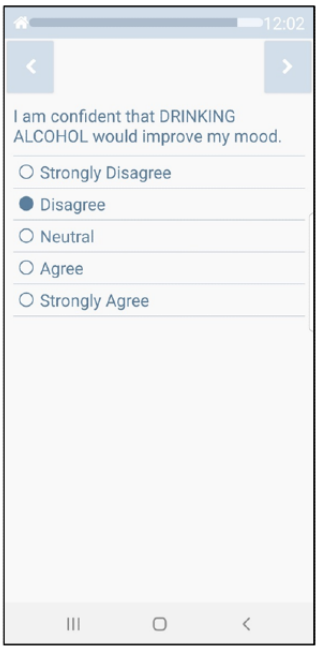
Daily Diary EMAs will be completed each day 30 min after the participant's self-reported wake time; questions will ask about the previous day (ie, "yesterday") and current (ie, "right now") experiences. Alcohol consumption will be assessed with the item "Did you drink any alcohol yesterday?" If the participant answers "yes," he/she will be prompted to indicate the number of standard drinks that were consumed. EMA reports have generally been seen as valid measures of drinking, even when participants are intoxicated [69,70]. See Figure 2. Additional items will assess sleeping arrangements from the prior night (example answer options: friend or family member's house or apartment, homeless shelter, jail, car, outside on the street), quality of sleep the previous night, social support and types of social interactions, stressors, other substance use, and substance abuse treatment attendance (see Table 3).

Table 3. Ecological momentary assessment (EMA) measures.

Type of EMA	Measure
Daily diary	• Sleeping arrangements • Social support and interactions • Treatment attendance • Current stressors and perceived stress • Alcohol consumption • Other substance use
Core/random/event sampling	• Affect/stress • Urge to drink • Alcohol availability • Social setting/location • Recent alcohol consumption • Expectancies • Abstinence motivation • Abstinence self-efficacy • New/ongoing stressful events[a] • Reasons for drinking[a] • Modified conflict tactics scale

[a]Drinking assessments only.

Figure 2. Smart-T alcohol example question.



Random Sampling

Participants will be prompted at random times to complete EMAs 4 times each day, scheduled to occur during the participant's normal waking hours. Participants will rate their affect by indicating the extent to which they agree or disagree with each of 13 statements at the moment: *I feel irritable, happy, content, angry, sad, worried, miserable,restless, stressed, hostile, calm, bored, anddepressed* (most items are from the circumplex model of affect [71]). In addition, participants will describe their current environment (eg, shelter, work, outside, or bar) and social setting (eg, alone, with others, or with others who are drinking). Alcohol urges (ie, "I have an urge to drink alcohol"; answer options range from strongly disagree to strongly agree), alcohol availability (ie, "Alcohol is available to me"; answer options range from not at all to easily available), drinking start/stop time, recent drinking, expectancies, motivation for abstinence, and abstinence self-efficacy will also be assessed during random sampling.

Event Sampling

Participants will be instructed to click the "I am About to Drink" or "I Just Drank" buttons if/when they have their first drink of the day. Drinking assessments will include all items from the random assessments and will query the reinforcing value of the drink or drinks and causes of the drinking episode.

Geocoding

The smartphones will be programmed to collect geolocation (ie, latitude, longitude) coordinates every 5 min.

Transdermal Alcohol Monitor

Biosensors can provide a continuous estimate of blood alcohol concentration (BAC) based on the concentration of alcohol in skin perspiration (Swift et al, unpublished data, August 1993) [72-78]. The device with the most extensive evaluation is the SCRAM bracelet (Alcohol Monitoring Systems Inc, Littleton, Colorado), which is worn on the ankle. SCRAM has an electrochemical sensor that samples the vapor from the skin every 30 min and stores readings for later retrieval. Peak transdermal alcohol concentration (TAC) is highly correlated with peak BAC, and self-reported volume of alcohol consumed is correlated with TAC estimates [79,80]. A recent project was able to detect 93.0% of self-reported heavy drinking episodes (≥5 drinks) using SCRAM data [81]. Data from the SCRAM are uploaded to SCRAMnet via USB cable and downloaded to a personal computer for analysis. The SCRAM is water resistant and cannot be removed without cutting the strap. To help address any stigma of wearing the device, participants will be provided with a laminated card that confirms their participation in the study and will be given a large (ie, bariatric) sock to wear over the equipment to make it less visible and more comfortable. At follow-up, participants will be asked about their experiences wearing the SCRAM.

Smartphone Hardware

A Samsung Galaxy J3 smartphone (or equivalent) will be loaned to each participant so that they may complete EMAs during the study. Participants enter data by touching their response on the screen (see Figure 2 for an example EMA item). Participants will be able to call (eg, if they have problems completing EMAs) and receive calls from research staff through the smartphone free of charge. The phone app encrypts data as they are collected and uploads data to the server multiple times per day.

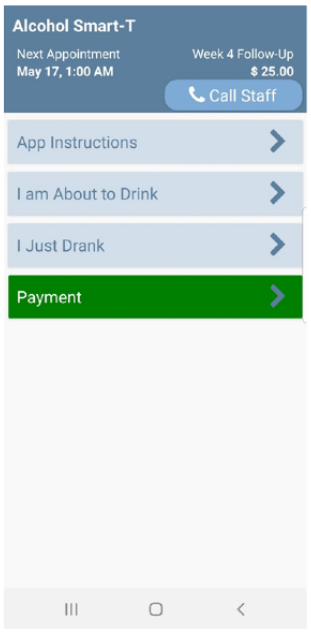
Smartphone Programming

The mobile health (mHealth) Shared Resource at the NCI Designated Stephenson Cancer Center will provide the programming services for this project. The mHealth resource employs a program manager, 2 research technicians, and 4 senior programmers who develop and maintain Web and mobile apps and relational databases. Apps are developed using the Insight platform, which consists of two components: a content management system (CMS) where researchers log in to CMS to create EMA/JITAI content and set EMA schedules, and a smartphone app shell. Once content is created, researchers transfer study materials into the smartphone app shell, greatly reducing the amount of time needed to create and deploy their smartphone app. The corresponding author is the Scientific Director of the mHealth Shared Resource.

Smartphone Training

Participants will watch a brief step-by-step video tutorial at the baseline visit that demonstrates how to use the app. This video will be loaded onto the app home screen (see “App Instructions” in Figure 3) so participants may view it at any time. The video will discuss how to complete EMAs and how to use the “Call Staff” and “Payment” button/options. We have achieved high EMA adherence rates (ie, 82%-87% of all EMAs completed) using similar protocols in previous samples of socioeconomically disadvantaged people (eg, homeless smokers and safety net hospital patients) [31,49].

Figure 3. Smart-T alcohol phase I home screen.



Compensation

Participants will be compensated with a US $25 gift card for completing the baseline assessment, US $25 for completing the 4-week assessment, up to US $25 in gift cards per week for completing EMAs (prorated based on percent completed), and US $25 for returning the phone and SCRAM in good condition at the end of the study. Specifically, those who complete 50% to 74% of EMA assessments each week will earn US $15 in gift cards, those who complete 75% to 89% of assessments will earn US $20 in gift cards, and those who complete 90% or more of their EMAs will earn US $25 in gift cards (payable at the 2-

and 4-week visits). Participants who complete less than 50% of the EMA prompts will not receive any compensation for the EMA component for that period. The phone shows the percent of EMAs completed. Overall, a participant can receive US $25 at the baseline visit, up to US $50 at the 2-week equipment check visit, and up to US $100 at the final visit.

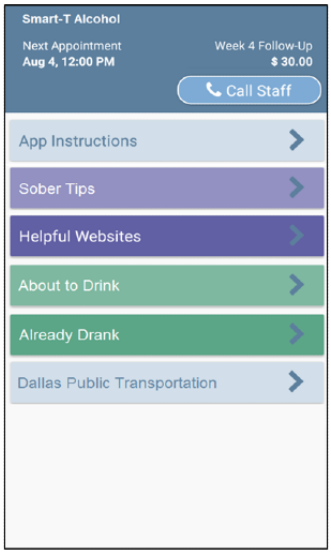
Phone Data Loss Prevention

To overcome potential loss of data if participants lose the study phone, phones will be programmed to connect to our secure server each day to upload encrypted data. This will minimize EMA data loss and allow researchers to monitor each participant's EMA completion rate and intervene when the rate is low. Importantly, EMA data are password-protected and encrypted on the study phone. Thus, study data are only accessible by the research team. If a phone is lost, it will be remotely wiped. We will provide one replacement phone if the participant has completed at least 50% of assessments for 1 week.

Phase III App

The phase III intervention app will have multiple components including (1) an on-demand "Tips" function/button, (2) a "Helpful Websites" function/button, (3) a "Call Staff" function/button, and (4) an algorithm that will use recent EMA responses and geolocation to assess current risk for alcohol use and automatically push relevant tailored messages to participants. The phone will record date/time when each of the components is accessed. See Figure 4 for the anticipated phase III home screen.

Figure 4. Anticipated Smart-T alcohol phase III home screen.



Sober Tips Function

Clicking this on-demand option will open a new window that will enable individuals to get useful tips related to "Benefits of Sobriety," "Motivational Messages," "Alcohol Refusal Skills," and "Managing Urges." The tips will be developed using strategies from previous motivational and skills-based interventions, such as the MAPIT [50], m.chat [82], and Smart-T [31] studies. For instance, when the "Managing Urges" tips option is clicked, participants will receive a suggestion on how to cope with their current urge to drink. This function will enable participants to access tailored messages at any time. Participants may view additional tips by clicking the "Next" button. Type and number of tips viewed will be recorded by the smartphone. Other topics for tips will be identified via examination of phase I participants' survey and EMA data.

Helpful Websites Function

Clicking this option will open a menu of useful websites (eg, Dallas public transit routes, Google maps, and online support groups such as Alcoholics Anonymous).

Call Staff Function

Clicking this option will connect participants to study staff if they encounter problems with the study phone and for scheduling/rescheduling follow-up appointments.

Risk Algorithm

The algorithm used to guide the just-in-time treatment messages will be similar to the algorithm that was developed for the Smart-T smoking cessation app [31]. Specifically, the algorithm will estimate risk for alcohol use using variables identified in phase I. In Smart-T, we attempted to develop risk algorithms that could predict smoking at 8, 12, and 24 hours before the lapse, but these algorithms were far less sensitive than the 4-hour lapse prediction algorithm. The resulting Smart-T algorithm combined six EMA variables (ie, urge, stress, cigarette availability, alcohol use, motivation to quit, and proximity to others smoking) to successfully predict 80% of all smoking lapses within 4 hours of lapse occurrence (false positive rate=17%) [31,32].

In phase III, smartphones will push tailored messages based on the momentary risk algorithm score at the end of each EMA. We anticipate that participant responses that indicate low risk for imminent alcohol use (eg, within the next 4 hours) will prompt delivery of level 1 messages. Level 1 messages will

primarily focus on increasing motivation for abstinence, avoiding people/places/things that may trigger alcohol use, benefits of sobriety, advice on ways to escape high-risk situations, and advice to seek support from others [50]. These messages will complement the treatment themes for those who are in an alcohol treatment program. Level 2 messages will be delivered at the end of EMAs if the algorithm determines that there is heightened risk for imminent (eg, within the next 4-8 hours) alcohol use. These messages will focus on in-the-moment distraction, reframing, immediate help-seeking, planning, and other tools to reduce craving. The highest rated indicator/trigger of alcohol use in that moment will be the topic of the level 2 tailored treatment messages. For example, if a participant reports low motivation for sobriety and average ratings on the other variables, they will receive a message that aims to boost motivation. An example may read: "You said that family was an important reason for staying sober. You're looking forward to a better life!" Likewise, if exposure to drinkers is an identified alcohol use trigger and a participant reports that he/she is near individuals who are consuming alcohol, he/she may receive a tailored suggestion on how to escape that high-risk situation, such as "You said that removing yourself from a situation was often helpful in managing cravings. Some people decide to get out of the situation, before they are tempted to drink." Participants will receive level 3 messages when they report recent drinking. Level 3 messages will focus on reframing the drinking episode as a learning experience and considering strategies for handling the situation differently in the future. We will draw from best-practice recommendations around message content and tone [83]. Our past interventions have contained hundreds of possible message combinations, depending on a person's baseline profile and current responses.

At the completion of phase II, we will have a working app that includes all components described above. The app will utilize EMA data (eg, location, time of day, urge to drink, affect [positive and negative], motivation, abstinence self-efficacy, and nearby presence of others who are drinking) to calculate risk and automatically intervene to reduce alcohol use in real time.

Statistical Analyses

Our assessment protocol is designed to capture diurnal patterns of experience and behavior within, between, and across days. Thus, data will have multiple time scales nested within individuals, ranging from every 5 min to monthly (ie, 288 geolocation assessments per day, 48 TAC readings per day, 5 prompted EMAs per day, participant-initiated drinking event EMAs, and monthly in-person assessments). Time and geolocation will be used as the variables to record risk and protective variables in calibrating one's risk. Traditional generalized linear mixed models; machine learning algorithms, such as elastic net penalized cox proportional hazards regression [84], or random forests [85], as well as spectral and dynamic modeling analyses, if feasible, will be used to identify predictors of study outcomes, model auto-regressive cyclical patterns, and capture intra- and interpersonal risk processes. We will divide the sample into training and testing datasets to validate the algorithm. Examples of planned analyses include (1) testing if alcohol urges or measures of affect predict daily drinking status, (2) testing if protective factors (eg, social support, positive interpersonal interactions, and time and location) predict alcohol use, and (3) testing if parameters of key variables (ie, intercept; slope, eg, increasing urges to drink over time; quadratic term; and volatility, eg, the symptom scatter or the ups and downs of urges over time) predict alcohol use.

Different levels of risk across individuals will also be included to examine how intraday risk gets intensified or ameliorated by personal trait-level variables (eg, sex, psychosocial resources, stress/adversity, and negative mood). We expect that analyses of SCT and ecological constructs (eg, affect, expectancies, self-efficacy, and proximity to drinking areas), gathered during random EMAs and breadcrumb trail geolocation, will identify patterns that predict drinking in near real time. In addition, EMA data will allow us to examine other important methodological questions such as (1) agreement between SCRAM, Timeline Follow-Back, and EMA reports of alcohol use and (2) the effect of episodic events (eg, exposure to violence or other stressors) on self-efficacy and mood and what impact that has on alcohol use. We will use SCRAM-detected alcohol use to explore the utility of SCRAM for validation of self-reported use events. We will use a macro developed by Barnett and colleagues [86] to interpret TAC data. We anticipate some missing data because of SCRAM bracelet malfunction, participant nonadherence, or participant attrition. In Barnett's work, data loss because of bracelet malfunction occurred on less than 5% of days of data collection, though we expect data loss to be somewhat higher in this study because of the nature of the population.

Questions similar to those listed in Table 4 will be used to assess the feasibility, acceptability, and usability of the phase III app. We will compare phase III participants (ie, EMAs and app features including tailored treatment messages) to phase I participants (ie, EMAs only) to examine the preliminary effectiveness of the app. Specifically, we will compare phase I and phase III participants' percent drinking days (PDD) and percent heavy drinking days (PHDD; ≥5 drinks for men, ≥4 drinks for women) using generalized linear models with an appropriate link function to accommodate outcome distributions. We will consider the study phase as the parameter of interest estimating the treatment effect, adjusting for relevant covariates (eg, gender, race, and baseline AUD severity). Exploratory analyses will examine the effect of specific types of treatment messages on intervention targets. For example, we will examine whether phase III participants' urges to drink are attenuated in EMAs that follow urge messages, and if postmessage reductions in urge are different/greater than that of phase I participants.

Table 4. Example acceptability items.

Question	Answer range
Overall, how helpful were the messages at the end of each assessment?	0=Not at all, 5=Extremely
Did the assessments and messages help you to make decisions that supported sobriety?	0=Definitely no, 5=Definitely yes
Overall, how helpful has the smartphone app been in helping you stay sober?	0=Not at all, 5=Extremely
How likely would you be to recommend the app to a friend?	0=Not at all, 5=Extremely

Results

The North Texas Regional Institutional Review Board approved the protocol as presented in this study in October 2018. The phase I smartphone app has been developed (see Figure 3), and data collection began in February 2019. Phase III data collection is expected to conclude in 2020. To date, 80 participants have consented to the study, and data analysis for phase I will begin in early 2020.

Discussion

Phase I

Phase I will advance previous research by identifying trends in behaviors, cognitions, and geolocation that predict subsequent drinking. This information will guide the development of JITAI in phase II that will be pilot-tested in phase III. We expect in-person assessments to show how psychosocial resources, stress/adversity, negative affect, and exposure to other drinkers and drinking locations affect alcohol use. We expect that analyses of SCT and ecological constructs (eg, affect, expectancies, self-efficacy, and proximity to drinking areas), gathered during EMAs and breadcrumb trail geolocation, will identify patterns that predict drinking in near real time. Findings from these analyses will identify additional targets for the intervention that will be developed in phase II.

Phase I will provide the foundation for one of the first smartphone interventions to be evaluated among adults who are homeless. As this population often lacks access to traditional intervention programs, the use of smartphone technology has tremendous potential to remove and attenuate barriers to service utilization for this at-risk population. Regardless of our initial intervention results, this work will provide valuable data on the array of risk and protective factors at multiple levels across multiple contexts affecting decision making and alcohol-use behavior in homeless adults.

Phases II and III

During phase II, we will create a “real time” drinking risk algorithm that can be used to deliver tailored treatment messages based upon current estimated risk of alcohol use. The integration of individual-level environmental exposure data alongside state-of-the-art EMA methodology sets this project apart from past studies. A strength of this project is that it does not rely solely on self-reported data but also embraces the “ecological” part of EMA by linking an individual’s behavior to the real-time environment in both time and location. Continuous “bread-crumb trail” geo-tracking provides a within-person control by documenting a person’s behavior with and without the presence of risk-promoting factors [87-89].

Treatment tailoring is most often done using participant characteristic or characteristics that are assessed at the baseline visit (eg, sex, and level of dependence). The proposed intervention will take this approach one step further by tailoring messages based on real-time risk for alcohol use. Phase III will test the initial efficacy of a smartphone app that assesses risk for alcohol use and automatically intervenes with tailored, theory-based treatment messages based on level of risk. The additive design of this project provides an analysis framework that will allow us to preliminarily examine the comparative effectiveness of the automated treatment messages triggered by our (phase II algorithm) relative to the calibration sample recruited in phase I. We expect that phase III participants will rate the app as helpful and useful and to report that it helped them to make decisions that were supportive of abstinence. We also anticipate preliminary evidence that phase III participants will demonstrate lower PDD and PHDD compared with phase I participants.

Limitations

Several study limitations warrant mention here. First, the intervention app that will be developed and tested during phase III may not be applicable to those with low literacy and cognitive impairment (eg, a 7th grade reading level is required to read and understand the smartphone-based assessments and intervention content). Second, only those who have access to a smartphone and access to electrical outlets to charge their phone will benefit from this type of intervention. It is important to note that research from our laboratory in two cities (ie, Dallas and Oklahoma City) has indicated that most adults experiencing homelessness possess phones with active cellular service. Finally, it is possible that EMA or SCRAM monitoring may have independent effects on drinking, especially given the frequency of assessment in this study. Some studies have found that self-monitoring can lead to changes in drinking, even without an “intended” intervention [90,91]. However, other studies have found only small, time-limited effects of frequent assessment, and in a study of college drinkers, frequent self-report and SCRAM monitoring did not strongly affect drinking behavior [92]. Nevertheless, we acknowledge that the design of this study will test the difference between assessment alone versus assessment + intervention messages and that potential therapeutic effect of monitoring needs to be studied in future studies.

Future Directions

This study will identify real-time antecedents of drinking among adults who are homeless, develop algorithms to predict risk of

alcohol use and tailored treatment messages, and provide pilot data on the efficacy of JITAI. Many people do not respond adequately to currently available treatment formats. This is particularly evident in underserved populations such as homeless adults, where treatment adherence tends to be poor. This project will gather valuable data on the risk and protective factors that affect drinking among adults experiencing homelessness. This information will be critical to developing other innovative treatments for this understudied and underserved population.

Acknowledgments

This work was supported by NIAAA grant R34AA024584 awarded to MB and SW. This work also was partially supported by the University of Oklahoma Health Sciences Center, Oklahoma Tobacco Settlement Endowment Trust (TSET) grant 092-016-0002, and NCI grant P30CA225520 awarded to the Stephenson Cancer Center, and the work used the mHealth Shared Resource.

Conflicts of Interest

MB is an inventor of the Insight mHealth Platform and receives royalties related to use of this platform.

References

1. Toro PA, Tompsett CJ, Lombardo S, Philippot P, Nachtergael H, Galand B, et al. Homelessness in Europe and the United States: a comparison of prevalence and public opinion. J Soc Issues 2007;63(3):505-524. [doi: 10.1111/j.1540-4560.2007.00521.x]
2. Barrow SM, Herman DB, Córdova P, Struening EL. Mortality among homeless shelter residents in New York City. Am J Public Health 1999 Apr;89(4):529-534. [doi: 10.2105/ajph.89.4.529] [Medline: 10191796]
3. Hwang SW, Wilkins R, Tjepkema M, O'Campo PJ, Dunn JR. Mortality among residents of shelters, rooming houses, and hotels in Canada: 11 year follow-up study. Br Med J 2009 Oct 26;339:b4036 [FREE Full text] [doi: 10.1136/bmj.b4036] [Medline: 19858533]
4. Weinreb L, Goldberg R, Perloff J. Health characteristics and medical service use patterns of sheltered homeless and low-income housed mothers. J Gen Intern Med 1998 Jun;13(6):389-397 [FREE Full text] [doi: 10.1046/j.1525-1497.1998.00119.x] [Medline: 9669568]
5. Kushel MB, Evans JL, Perry S, Robertson MJ, Moss AR. No door to lock: victimization among homeless and marginally housed persons. Arch Intern Med 2003 Nov 10;163(20):2492-2499. [doi: 10.1001/archinte.163.20.2492] [Medline: 14609786]
6. Morrison DS. Homelessness as an independent risk factor for mortality: results from a retrospective cohort study. Int J Epidemiol 2009 Jun;38(3):877-883. [doi: 10.1093/ije/dyp160] [Medline: 19304988]
7. Hasin DS, Stinson FS, Ogburn E, Grant BF. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry 2007 Jul;64(7):830-842. [doi: 10.1001/archpsyc.64.7.830] [Medline: 17606817]
8. North CS, Eyrich-Garg KM, Pollio DE, Thirthalli J. A prospective study of substance use and housing stability in a homeless population. Soc Psychiatry Psychiatr Epidemiol 2010 Nov;45(11):1055-1062. [doi: 10.1007/s00127-009-0144-z] [Medline: 19816646]
9. Koegel P, Burnam MA, Farr RK. The prevalence of specific psychiatric disorders among homeless individuals in the inner city of Los Angeles. Arch Gen Psychiatry 1988 Dec;45(12):1085-1092. [doi: 10.1001/archpsyc.1988.01800360033005] [Medline: 2461690]
10. Bassuk EL, Buckner JC, Perloff JN, Bassuk SS. Prevalence of mental health and substance use disorders among homeless and low-income housed mothers. Am J Psychiatry 1998 Nov;155(11):1561-1564. [doi: 10.1176/ajp.155.11.1561] [Medline: 9812118]
11. Breakey WR, Fischer PJ, Kramer M, Nestadt G, Romanoski AJ, Ross A, et al. Health and mental health problems of homeless men and women in Baltimore. J Am Med Assoc 1989 Sep 8;262(10):1352-1357. [Medline: 2761036]
12. Baggett TP, Chang Y, Singer DE, Porneala BC, Gaeta JM, O'Connell JJ, et al. Tobacco-, alcohol-, and drug-attributable deaths and their contribution to mortality disparities in a cohort of homeless adults in Boston. Am J Public Health 2015 Jun;105(6):1189-1197. [doi: 10.2105/AJPH.2014.302248] [Medline: 25521869]
13. Burt M, Aron L, Douglas T, Valente J, Lee E, Iwen B. Homelessness: programs and the people they serve. In: Findings of the National Survey of Homeless Assistance Providers. Washington, DC: Department of Housing and Urban Development; Dec 1999.
14. Wenzel SL, Burnam MA, Koegel P, Morton SC, Miu A, Jinnett KJ, et al. Access to inpatient or residential substance abuse treatment among homeless adults with alcohol or other drug use disorders. Med Care 2001 Nov;39(11):1158-1169. [doi: 10.1097/00005650-200111000-00003] [Medline: 11606870]
15. Kertesz SG, Larson MJ, Cheng DM, Tucker JA, Winter M, Mullins A, et al. Need and non-need factors associated with addiction treatment utilization in a cohort of homeless and housed urban poor. Med Care 2006 Mar;44(3):225-233. [doi: 10.1097/01.mlr.0000199649.19464.8f] [Medline: 16501393]

16. Gonzalez G, Rosenheck RA. Outcomes and service use among homeless persons with serious mental illness and substance abuse. Psychiatr Serv 2002 Apr;53(4):437-446. [doi: 10.1176/appi.ps.53.4.437] [Medline: 11919357]
17. Argeriou M, McCarty D. The use of shelters as substance abuse stabilization sites. J Ment Health Adm 1993;20(2):126-137. [doi: 10.1007/bf02519237] [Medline: 10171528]
18. Bradford DW, Gaynes BN, Kim MM, Kaufman JS, Weinberger M. Can shelter-based interventions improve treatment engagement in homeless individuals with psychiatric and/or substance misuse disorders?: a randomized controlled trial. Med Care 2005 Aug;43(8):763-768. [doi: 10.1097/01.mlr.0000170402.35730.ea] [Medline: 16034289]
19. Scott-Lennox J, Rose R, Bohlig A, Lennox R. The impact of women's family status on completion of substance abuse treatment. J Behav Health Serv Res 2000 Nov;27(4):366-379. [doi: 10.1007/bf02287819] [Medline: 11070631]
20. Schonfeld L, Dupree LW, Dickson-Euhrmann E, Royer CM, McDermott CH, Rosansky JS, et al. Cognitive-behavioral treatment of older veterans with substance abuse problems. J Geriatr Psychiatry Neurol 2000;13(3):124-129. [doi: 10.1177/089198870001300305] [Medline: 11001134]
21. Orwin RG, Garrison-Mogren R, Jacobs ML, Sonnefeld LJ. Retention of homeless clients in substance abuse treatment. Findings from the National Institute on Alcohol Abuse and Alcoholism Cooperative Agreement Program. J Subst Abuse Treat 1999;17(1-2):45-66. [doi: 10.1016/s0740-5472(98)00046-4] [Medline: 10435252]
22. Shiffman S, Hufford M, Hickcox M, Paty JA, Gnys M, Kassel JD. Remember that? A comparison of real-time versus retrospective recall of smoking lapses. J Consult Clin Psychol 1997 Apr;65(2):292-300. [doi: 10.1037/0022-006x.65.2.292.a] [Medline: 9086693]
23. Stone AA, Schwartz JE, Neale JM, Shiffman S, Marco CA, Hickcox M, et al. A comparison of coping assessed by ecological momentary assessment and retrospective recall. J Pers Soc Psychol 1998 Jun;74(6):1670-1680. [doi: 10.1037//0022-3514.74.6.1670] [Medline: 9654765]
24. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. Br Med J 2002 May 18;324(7347):1193-1194 [FREE Full text] [doi: 10.1136/bmj.324.7347.1193] [Medline: 12016186]
25. Kirchner TR, Shiffman S. Ecological momentary assessment. In: MacKillop J, editor. The Wiley-Blackwell Handbook of Addiction Psychopharmacology. Hoboken, NJ: Wiley-blackwell; 2013.
26. Morgenstern J, Kuerbis A, Muench F. Ecological momentary assessment and alcohol use disorder treatment. Alcohol Res 2014;36(1):101-109 [FREE Full text] [Medline: 26259004]
27. Kirchner TR, Shiffman S. Spatio-temporal determinants of mental health and well-being: advances in geographically-explicit ecological momentary assessment (GEMA). Soc Psychiatry Psychiatr Epidemiol 2016 Sep;51(9):1211-1223 [FREE Full text] [doi: 10.1007/s00127-016-1277-5] [Medline: 27558710]
28. Watkins KL, Regan SD, Nguyen N, Businelle MS, Kendzor DE, Lam C, et al. Advancing cessation research by integrating EMA and geospatial methodologies: associations between tobacco retail outlets and real-time smoking urges during a quit attempt. Nicotine Tob Res 2014 May;16(Suppl 2):S93-101 [FREE Full text] [doi: 10.1093/ntr/ntt135] [Medline: 24057995]
29. Shiffman S, Balabanis MH, Paty JA, Engberg J, Gwaltney CJ, Liu KS, et al. Dynamic effects of self-efficacy on smoking lapse and relapse. Health Psychol 2000 Jul;19(4):315-323. [doi: 10.1037//0278-6133.19.4.315] [Medline: 10907649]
30. Reitzel LR, Kendzor DE, Nguyen N, Regan SD, Okuyemi KS, Castro Y, et al. Shelter proximity and affect among homeless smokers making a quit attempt. Am J Health Behav 2014 Mar;38(2):161-169 [FREE Full text] [doi: 10.5993/AJHB.38.2.1] [Medline: 24629545]
31. Businelle MS, Ma P, Kendzor DE, Frank SG, Vidrine DJ, Wetter DW. An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness. J Med Internet Res 2016 Dec 12;18(12):e321 [FREE Full text] [doi: 10.2196/jmir.6058] [Medline: 27956375]
32. Businelle MS, Ma P, Kendzor DE, Frank SG, Wetter DW, Vidrine DJ. Using intensive longitudinal data collected via mobile phone to detect imminent lapse in smokers undergoing a scheduled quit attempt. J Med Internet Res 2016 Oct 17;18(10):e275 [FREE Full text] [doi: 10.2196/jmir.6307] [Medline: 27751985]
33. Businelle MS, Ma P, Kendzor DE, Reitzel LR, Chen M, Lam CY, et al. Predicting quit attempts among homeless smokers seeking cessation treatment: an ecological momentary assessment study. Nicotine Tob Res 2014 Oct;16(10):1371-1378 [FREE Full text] [doi: 10.1093/ntr/ntu088] [Medline: 24893602]
34. Kuerbis A, Armeli S, Muench F, Morgenstern J. Motivation and self-efficacy in the context of moderated drinking: global self-report and ecological momentary assessment. Psychol Addict Behav 2013 Dec;27(4):934-943 [FREE Full text] [doi: 10.1037/a0031194] [Medline: 23276318]
35. Freedman MJ, Lester KM, McNamara C, Milby JB, Schumacher JE. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. J Subst Abuse Treat 2006 Mar;30(2):105-111. [doi: 10.1016/j.jsat.2005.10.005] [Medline: 16490673]
36. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
37. Kirchner TR, Shiffman S, Wileyto EP. Relapse dynamics during smoking cessation: recurrent abstinence violation effects and lapse-relapse progression. J Abnorm Psychol 2012 Feb;121(1):187-197 [FREE Full text] [doi: 10.1037/a0024451] [Medline: 21787035]

38. Bronfenbrenner U. The Ecology Of Human Development: Experiments By Nature And Design. Cambridge, MA: Harvard University Press; 1979.
39. Glass TA, McAtee MJ. Behavioral science at the crossroads in public health: extending horizons, envisioning the future. Soc Sci Med 2006 Apr;62(7):1650-1671. [doi: 10.1016/j.socscimed.2005.08.044] [Medline: 16198467]
40. Bewick BM, Trusler K, Barkham M, Hill AJ, Cahill J, Mulhern B. The effectiveness of web-based interventions designed to decrease alcohol consumption--a systematic review. Prev Med 2008 Jul;47(1):17-26. [doi: 10.1016/j.ypmed.2008.01.005] [Medline: 18302970]
41. Carver CS, Scheier M. On the Self-Regulation of Behavior. Cambridge, UK: Cambridge University Press; 1998.
42. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
43. Weitzel JA, Bernhardt JM, Usdan S, Mays D, Glanz K. Using wireless handheld computers and tailored text messaging to reduce negative consequences of drinking alcohol. J Stud Alcohol Drugs 2007 Jul;68(4):534-537. [doi: 10.15288/jsad.2007.68.534] [Medline: 17568957]
44. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A context-sensing mobile phone app (Q Sense) for smoking cessation: a mixed-methods study. JMIR Mhealth Uhealth 2016 Sep 16;4(3):e106 [FREE Full text] [doi: 10.2196/mhealth.5787] [Medline: 27637405]
45. Vahabzadeh M, Mezghanni M, Jia-Ling L, Epstein D, Preston K. PGIS: Electronic Diary Data Integration With GPS Data Initial Application in Substance-Abuse Patients. In: Proceedings of the 2010 IEEE 23rd International Symposium on Computer-Based Medical Systems. 2010 Presented at: CBMS'10; October 12-15, 2010; Perth, WA, Australia. [doi: 10.1109/cbms.2010.6042691]
46. Dulin PL, Gonzalez VM, Campbell K. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: usability and early outcomes. Subst Abus 2014;35(2):168-175 [FREE Full text] [doi: 10.1080/08897077.2013.821437] [Medline: 24821354]
47. Agyapong VI, McLoughlin DM, Farren CK. Six-months outcomes of a randomised trial of supportive text messaging for depression and comorbid alcohol use disorder. J Affect Disord 2013 Oct;151(1):100-104. [doi: 10.1016/j.jad.2013.05.058] [Medline: 23800443]
48. Gonzalez VM, Dulin PL. Comparison of a smartphone app for alcohol use disorders with an internet-based intervention plus bibliotherapy: a pilot study. J Consult Clin Psychol 2015 Apr;83(2):335-345 [FREE Full text] [doi: 10.1037/a0038620] [Medline: 25622202]
49. Businelle MS, Kendzor DE, Kesh A, Cuate EL, Poonawalla IB, Reitzel LR, et al. Small financial incentives increase smoking cessation in homeless smokers: a pilot study. Addict Behav 2014 Mar;39(3):717-720. [doi: 10.1016/j.addbeh.2013.11.017] [Medline: 24321696]
50. Walters ST, Ondersma SJ, Ingersoll KS, Rodriguez M, Lerch J, Rossheim ME, et al. MAPIT: development of a web-based intervention targeting substance abuse treatment in the criminal justice system. J Subst Abuse Treat 2014 Jan;46(1):60-65 [FREE Full text] [doi: 10.1016/j.jsat.2013.07.003] [Medline: 23954392]
51. Businelle MS, Poonawalla IB, Kendzor DE, Rios DM, Cuate EL, Savoy EJ, et al. Smoking policy change at a homeless shelter: attitudes and effects. Addict Behav 2015 Jan;40:51-56. [doi: 10.1016/j.addbeh.2014.08.013] [Medline: 25222848]
52. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
53. Adler N, Stewart J. MacArthur SES & Health Network. 2007 Mar. The MacArthur Scale of Subjective Social Status URL: http://www.macses.ucsf.edu/Research/Psychosocial/subjective.php [accessed 2011-03-01]
54. Tsemberis S, McHugo G, Williams V, Hanrahan P, Stefancic A. Measuring homelessness and residential stability: The residential time-line follow-back inventory. J Commun Psychol 2007;35(1):29-42. [doi: 10.1002/jcop.20132]
55. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
56. Centers for Disease Control and Prevention. Atlanta, Georgia: US Department of Health and Human Services Behavioral Risk Factor Surveillance System: Survey Data & Documentation URL: https://www.cdc.gov/brfss/data_documentation/index.htm [accessed 2020-02-17]
57. Zajacova A, Dowd JB. Reliability of self-rated health in US adults. Am J Epidemiol 2011 Oct 15;174(8):977-983 [FREE Full text] [doi: 10.1093/aje/kwr204] [Medline: 21890836]
58. Centers for Disease Control and Prevention. Atlanta, GA: US Department of Health and Human Services; 2011. Behavioral Risk Factor Surveillance System: Questionnaire URL: https://www.cdc.gov/brfss/questionnaires/index.htm [accessed 2020-02-17]
59. Sobell LC, Ontario Addiction Research Foundation, Sobell MB. Timeline Followback user's guide: a calendar method for assessing alcohol and drug use. In: Timeline Followback Users Guide. Ontario: Addiction Research Foundation; 1996.
60. Miller WR, Tonigan JS, Longabaugh R. National Institute on Alcohol Abuse and Alcoholism. Rockville, MD: US Department of Health and Human Services; 1995. The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing

Adverse Consequences of Alcohol Abuse. Test Manual. (Project MATCH Monograph No. 4) URL: https://pubs.niaaa.nih.gov/publications/projectmatch/match04.pdf [accessed 2020-02-17]
61. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 1997 Aug 15;277(5328):918-924 [FREE Full text] [doi: 10.1126/science.277.5328.918] [Medline: 9252316]
62. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
63. Jaffee KD, Liu GC, Canty-Mitchell J, Qi RA, Austin J, Swigonski N. Race, urban community stressors, and behavioral and emotional problems of children with special health care needs. Psychiatr Serv 2005 Jan;56(1):63-69. [doi: 10.1176/appi.ps.56.1.63] [Medline: 15637194]
64. Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). Arch Intern Med 1999;159(15):1701-1704. [doi: 10.1001/archinte.159.15.1701] [Medline: 10448771]
65. Brookings J, Bolton B. Confirmatory factor analysis of the Interpersonal Support Evaluation List. Am J Community Psychol 1988 Feb;16(1):137-147. [doi: 10.1007/bf00906076] [Medline: 3369379]
66. Carver CS. You want to measure coping but your protocol's too long: consider the brief COPE. Int J Behav Med 1997;4(1):92-100. [doi: 10.1207/s15327558ijbm0401_6] [Medline: 16250744]
67. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
68. Wetter DW, McClure JB, Cofta-Woerpel L, Costello TJ, Reitzel LR, Businelle MS, et al. A randomized clinical trial of a palmtop computer-delivered treatment for smoking relapse prevention among women. Psychol Addict Behav 2011 Jun;25(2):365-371 [FREE Full text] [doi: 10.1037/a0022797] [Medline: 21500879]
69. Wray TB, Merrill JE, Monti PM. Using ecological momentary assessment (EMA) to assess situation-level predictors of alcohol use and alcohol-related consequences. Alcohol Res 2014;36(1):19-27 [FREE Full text] [Medline: 26258997]
70. Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 2009 Dec;21(4):486-497 [FREE Full text] [doi: 10.1037/a0017074] [Medline: 19947783]
71. Shiffman S, Kirchner TR. Cigarette-by-cigarette satisfaction during ad libitum smoking. J Abnorm Psychol 2009 May;118(2):348-359 [FREE Full text] [doi: 10.1037/a0015620] [Medline: 19413409]
72. Swift R. Direct measurement of alcohol and its metabolites. Addiction 2003 Dec;98(Suppl 2):73-80. [doi: 10.1046/j.1359-6357.2003.00605.x] [Medline: 14984244]
73. Phillips M, Greenberg J, Andrzejewski J. Evaluation of the Alcopatch, a transdermal dosimeter for monitoring alcohol consumption. Alcohol Clin Exp Res 1995 Dec;19(6):1547-1549. [doi: 10.1111/j.1530-0277.1995.tb01022.x] [Medline: 8749825]
74. Swift RM. Transdermal measurement of alcohol consumption. Addiction 1993 Aug;88(8):1037-1039. [doi: 10.1111/j.1360-0443.1993.tb02122.x] [Medline: 8401157]
75. Swift RM, Martin CS, Swette L, LaConti A, Kackley N. Studies on a wearable, electronic, transdermal alcohol sensor. Alcohol Clin Exp Res 1992 Aug;16(4):721-725. [doi: 10.1111/j.1530-0277.1992.tb00668.x] [Medline: 1530135]
76. Leffingwell TR, Cooney NJ, Murphy JG, Luczak S, Rosen G, Dougherty DM, et al. Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors. Alcohol Clin Exp Res 2013 Jan;37(1):16-22 [FREE Full text] [doi: 10.1111/j.1530-0277.2012.01869.x] [Medline: 22823467]
77. Hawthorne JS, Wojcik MH. Transdermal alcohol measurement: a review of the literature. Can Soc Forensic Sci J 2006;39(2):65-71. [doi: 10.1080/00085030.2006.10757138]
78. Litten RZ, Bradley AM, Moss HB. Alcohol biomarkers in applied settings: recent advances and future research opportunities. Alcohol Clin Exp Res 2010 Jun;34(6):955-967. [doi: 10.1111/j.1530-0277.2010.01170.x] [Medline: 20374219]
79. Sakai JT, Mikulich-Gilbertson SK, Long RJ, Crowley TJ. Validity of transdermal alcohol monitoring: fixed and self-regulated dosing. Alcohol Clin Exp Res 2006 Jan;30(1):26-33. [doi: 10.1111/j.1530-0277.2006.00004.x] [Medline: 16433729]
80. Dougherty DM, Charles NE, Acheson A, John S, Furr RM, Hill-Kapturczak N. Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. Exp Clin Psychopharmacol 2012 Oct;20(5):373-381 [FREE Full text] [doi: 10.1037/a0029021] [Medline: 22708608]
81. Barnett NP, Meade EB, Glynn TR. Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. Exp Clin Psychopharmacol 2014 Feb;22(1):86-96 [FREE Full text] [doi: 10.1037/a0034821] [Medline: 24490713]
82. Walters ST, Spence-Almaguer E, Hill W, Abraham S. Integrating health coaching and technology with vulnerable clients. Soc Work Today 2015;15(6) [FREE Full text]
83. Walters ST. Motivational Interviewing Network of Trainers (MINT). Reno, NV: Mountain Plains Addiction Technology Transfer Center, University of Nevada, Reno; 2019. A Guide to Using Text Messages to Improve Substance Use Treatment Outcomes URL: http://motivationalinterviewing.org/sites/default/files/walters_2019_guide_to_using_text_messages.pdf [accessed 2020-02-17]
84. Suchting R, Hébert ET, Ma P, Kendzor DE, Businelle MS. Using elastic net penalized cox proportional hazards regression to identify predictors of imminent smoking lapse. Nicotine Tob Res 2019 Jan 4;21(2):173-179. [doi: 10.1093/ntr/ntx201] [Medline: 29059349]

85. Breiman L. Random forests. Mach Learn 2001;45(1):5-32 [FREE Full text] [doi: 10.1023/A:1010933404324]
86. Barnett N, Souza T, Rosen I, Luczak S, Glynn T, Swift R. Brown University. 2015. Transdermal Alcohol Sensor Data Macro (Version 1.3) Software URL: https://www.brown.edu/academics/public-health/research/alcohol-addiction-studies/tasmac/ [accessed 2020-02-17]
87. Kirchner TR, Cantrell J, Anesetti-Rothermel A, Ganz O, Vallone DM, Abrams DB. Geospatial exposure to point-of-sale tobacco: real-time craving and smoking-cessation outcomes. Am J Prev Med 2013 Oct;45(4):379-385 [FREE Full text] [doi: 10.1016/j.amepre.2013.05.016] [Medline: 24050412]
88. Kirchner T, Gao H, Anesetti-Rothermel A, Carlos H, House B. Longitudinal Human Mobility and Real-time Access to a National Density Surface of Retail Outlets. In: Proceedings of ACM Urban Computing 2014. 2014 Presented at: UrbComp'14; August 24, 2014; New York, NY.
89. Kirchner TR, Cantrell J, Anesetti-Rothermel A, Pearson J, Cha S, Kreslake J, et al. Individual Mobility Patterns and Real-Time Geo-Spatial Exposure to Point-of-sale Tobacco Marketing. In: Proceedings of the conference on Wireless Health. New York, NY: ACM; 2012 Presented at: WH'12; October 23 - 25, 2012; San Diego, California. [doi: 10.1145/2448096.2448104]
90. Maisto SA, Clifford PR, Davis CM. Alcohol treatment research assessment exposure subject reactivity effects: part II. Treatment engagement and involvement. J Stud Alcohol Drugs 2007 Jul;68(4):529-533. [doi: 10.15288/jsad.2007.68.529] [Medline: 17568956]
91. Walters ST, Vader AM, Harris TR, Jouriles EN. Reactivity to alcohol assessment measures: an experimental test. Addiction 2009 Aug;104(8):1305-1310 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02632.x] [Medline: 19624323]
92. Luczak SE, Rosen IG, Wall TL. Development of a real-time repeated-measures assessment protocol to capture change over the course of a drinking episode. Alcohol Alcohol 2015 Mar;50(2):180-187 [FREE Full text] [doi: 10.1093/alcalc/agu100] [Medline: 25568142]

Abbreviations

AUD: alcohol use disorder
BAC: blood alcohol concentration
CMS: content management system
EMA: ecological momentary assessment
JITAI: just-in-time adaptive intervention
mHealth: mobile health
PDD: percent drinking days
PHDD: percent heavy drinking days
QDS: Questionnaire Development System
SCRAM: Secure Continuous Remote Alcohol Monitor
SCT: social cognitive theory
SUD: substance use disorder
TAC: transdermal alcohol concentration

Edited by G Eysenbach; submitted 23.07.19; peer-reviewed by R Vilardaga, M Raven; comments to author 02.12.19; revised version received 13.01.20; accepted 27.01.20; published 16.04.20.

Please cite as:
Businelle MS, Walters ST, Mun EY, Kirchner TR, Hébert ET, Li X
Reducing Drinking Among People Experiencing Homelessness: Protocol for the Development and Testing of a Just-in-Time Adaptive Intervention
JMIR Res Protoc 2020;9(4):e15610
URL: http://www.researchprotocols.org/2020/4/e15610/
doi:10.2196/15610
PMID:32297874

Protocol

Reducing Burnout and Promoting Health and Wellness Among Medical Students, Residents, and Physicians in Alberta: Protocol for a Cross-Sectional Questionnaire Study

Esther Kim[1], MD; Robert Mallett[1], MD; Marianne Hrabok[1,2], PhD; Yajing Alicia Yang[3], PMP, MSc; Chantal Moreau[1], MD, FRCPC; Izu Nwachukwu[2], MD, FRCPC; Maryana Kravtsenyuk[1], MSc, MD, FRCPC; Adam Abba-Aji[1], MD, FRCPC; Daniel Li[1], MD, FRCPC; Vincent I O Agyapong[1], MD, PhD, FRCPC

[1]Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

[2]Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3]Alberta Health Services, Edmonton, AB, Canada

Corresponding Author:
Vincent I O Agyapong, MD, PhD, FRCPC
Department of Psychiatry
Faculty of Medicine and Dentistry
University of Alberta
1E1 Walter Mackenzie Health Sciences Centre
8440 112 St NW
Edmonton, AB, T6G 2B7
Canada
Phone: 1 780 215 7771
Email: agyapong@ualberta.ca

Abstract

Background: Burnout is an increasingly common and insidious phenomenon experienced by workers in many different fields, although it is of particular concern among physicians and trainees due to the nature of their work. It is estimated that one-third of practicing physicians will experience burnout during their career, and this rate is expected to continue to increase. Burnout has significant implications, as it has been identified as a contributor to increased medical errors, decreased patient satisfaction, substance use, workforce attrition, and suicide.

Objective: This study will evaluate the prevalence and impact of burnout on physicians, residents, and medical students in Alberta.

Methods: Quantitative and qualitative data collected through self-administered, anonymous, online questionnaires will be used in this cross-sectional provincial study design. Data collection tools were developed based on published literature and questions from previously validated instruments. The tools capture relevant demographic information, mental health status, and rates of burnout, as well as factors contributing to both burnout and resilience among respondents. We anticipate a sample size of 777 medical students, 959 residents, and 1961 physicians to represent the respective ratios of trainees and practicing physicians in the province of Alberta.

Results: Study recruitment will begin in September 2020, with 4 weeks of data collection. The results of this study are anticipated within 12 months from the end of data collection. It is expected that the results will provide an overview of the prevalence of burnout among those training and working in medicine in Alberta, identify contributors to burnout, and help develop interventions aimed at reducing burnout.

Conclusions: This study's aim is to examine burnout prevalence and contributing factors among medical trainees and physicians in Alberta. It is expected that the results will identify and examine individual and organizational practices that contribute to burnout and help develop strategies and interventions focused on mitigating burnout and its sequelae.

International Registered Report Identifier (IRRID): PRR1-10.2196/16285

(JMIR Res Protoc 2020;9(4):e16285) doi:10.2196/16285

KEYWORDS

burnout, psychological; wellness; resilience; burnout measures; burnout interventions; delivery of health care; medicine; work life balance

Introduction

Burnout is a complex phenomenon that is characterized by three domains: emotional exhaustion, depersonalization, and reduced personal accomplishment [1]. Emotional exhaustion refers to excessive emotional fatigue from overwhelming demands in the workplace; depersonalization refers to the impersonal feelings or indifference experienced towards patients; and reduced personal accomplishment refers to deficits in one's conviction that they are competent and successful in their work [1]. The term burnout was coined in the 1960s [2] to describe the psychological and emotional distress experienced by health care staff working at a clinic for the underprivileged. Today, the term reflects an individual's negative response to the demands of the workplace in the setting of chronic affective and interpersonal stressors [3-6]. In the medical profession, it is felt that chronic exposure to stress is the primary cause of emotional exhaustion, while depersonalization occurs as physicians begin to treat patients indifferently and develop some degree of pessimism towards their colleagues and profession. A lack of personal achievement is thought to occur when practitioners withdraw from their responsibilities at work [7]. This study protocol aims to provide an overview of the prevalence of burnout among medical trainees and physicians within Alberta and to identify systemic and personal factors contributing to burnout in order to develop potentially mitigating interventions.

Among the general working population, the prevalence of burnout is suggested to range from 7% to 25%, with discrepancy among rates secondary to the use of different scales and cutoff levels in population-based studies [8-11]. The majority of burnout studies have focused on specific subgroups within the context of human services occupations, including but not limited to education, health care, and support services. While burnout is not exclusive to physicians, the profession's increasing workload, challenging work environment, increasing documentation demands (including the emergence of electronic health records), threats of litigation, and challenges to work-life balance render physicians particularly vulnerable to burnout [12]. Burnout affects physicians at all stages of their careers. Research suggests that burnout takes root during medical school and continues into residency and practice [13]. Romani and Ashkar [7] reported that up to 50% of medical students, 76% of surgical and internal medicine residents, and 45% of practicing physicians in the United States reported at least one symptom of burnout. These frequencies, when compared to the general working population, have raised concerns about burnout and its effects on the medical profession.

Medical students with high burnout scores were more likely to engage in unprofessional behaviors, such as cheating and plagiarism, which may undermine future professionalism (eg, managing conflicts of interest, reporting incompetent colleagues, adhering to appropriate prescription practices) [14]. Additional research has suggested that higher burnout scores are associated with lower empathy scores, less altruistic views of medicine as a career, and consideration of leaving the medical profession altogether [13,15-17]. A national study of medical students in the United States reported that increased risk of alcohol abuse was independently associated with burnout [18]. Burnout in students was also a predictor of suicidal ideation, which is particularly worrisome when considered in combination with alcohol abuse [19]. Dyrbye et al [13] found that depression and suicidal ideation were most pronounced during medical school and diminished as individuals' careers progressed through training into practice.

In first-year internal medicine residents, higher burnout scores were associated with lower reported quality of life and education with more negative patient-doctor relationships, negative interactions with other health care professionals, and perceived increase in medical errors with poorer patient safety [20,21]. Burnout, depersonalization, and fatigue were highest during residency, with fatigue being an independent factor contributing to errors, injuries, and motor vehicle collisions [13]. In a US national study involving internal medicine residents, higher burnout scores, specifically high emotional exhaustion, were associated with lower scores on a standardized national exam [22]. Suicidal thoughts were also more prevalent in residents with burnout than those without [23]. It is suggested that the likelihood of burnout is highest within the first year of residency but can persist throughout the duration of residency [24].

Overall, one-third of all practicing physicians are expected to experience burnout during their career, and the rates of burnout appear to be worsening over time [25]. It appears that rates of burnout among practicing physicians are relatively consistent across geographical regions, with one-third of physicians in Yemen, Qatar, and Saudi Arabia reporting burnout, similar to self-reported rates in the United States [7]. Burnout among physicians increases medical errors, negatively impacts patient safety, and leads to lower patient satisfaction [2,26]. Alcohol abuse was independently associated with burnout among US physicians and surgeons [27,28]. Grinspoon [12] questioned the role of burnout in the roughly 400 physician suicides that occur each year in the United States and noted the increasing numbers of physicians leaving the profession mid-career.

The phenomenon of burnout has yet to be defined by specific criteria. Until now, burnout has typically been measured using the Maslach Burnout Inventory, a 22-question, self-reported, validated survey that assesses the key domains of emotional exhaustion, depersonalization, and personal accomplishment on a 7-point Likert scale [1]. Alternative scales have also been developed and used in physician and medical trainee populations to further characterize the heterogeneity of burnout [29-33]. For example, the Oldenburg Burnout Inventory and Utrechtse Burnout Scale distinguish burnout domains in the context of job demands and job resources [33,34]. The Copenhagen Burnout Inventory describes burnout domains in relation to personal, work, and client-related factors [35]. These alternative

scales reflect the multifaceted interplay of diverse factors that contribute to burnout among physicians and medical trainees [29-33].

Using various scales in qualitative and quantitative studies, evidence has shown that burnout is associated with a variety of individual, occupational, and organizational factors. Based on the 5-dimensional model of personality coined by Goldberg in 1990, neuroticism has been shown to potentially predispose individuals to developing burnout through maladaptive coping mechanisms [36]. Longitudinal data suggest associations exist between neuroticism and emotional exhaustion among physicians and that these physicians report less satisfaction with medicine as a career [37]. Psychosocial stressors outside of medicine and the workplace such as illness, family-related conflict, or financial concerns can also increase trainees' vulnerability to burnout [24,38,39]. For instance, residents with significant amounts of educational debt were more likely to have higher rates of burnout [22].

Burnout rates are not only affected by work-related factors such as area of practice, hours worked per week, lower autonomy, and number of call shifts per week but also non work-related factors such as age, gender, and number of children [7,40,41]. Among Hungarian general practitioners and residents, younger age, male gender, fewer years of experience, and increased number of dependents such as children were correlated with higher rates of burnout [40]. Such results were postulated to be secondary to less work experience leading to increased stress in the workplace as well as difficulty with balancing the demands of home and work life, leading to higher levels of emotional exhaustion and perpetuating depersonalization [40].

Unfortunately, there is a dearth of data regarding interventions that have proven beneficial in reducing burnout rates, including stress management courses, mindfulness, brief exercise, and short-term counselling. In interventions that have been implemented, there is insufficient evidence to suggest that they have a meaningful impact past the intervention period [7]. Furthermore, interventions aimed at reducing burnout must cater specifically to each stage of practice; it is unlikely that a single solution will be uniformly effective for medical school training, residency, and independent practice [7].

Within Canada, physician health and wellness initiatives have been evolving, with the development of professional support networks, wellness curriculums, adoption of health as a component of core CanMEDS competencies, and risk-management strategies adapted from organizational and occupational systems [42,43]. However, despite targeted strategies, such as mindfulness-based therapy, resilience training, and access to professional services, the 2018 Canadian Medical Association National Physician Health Survey suggests ongoing high rates of burnout (30%) and positive screening for depression (34%) among physicians and residents [44]. Additionally, there are insufficient data regarding burnout rates within each province in Canada. This poses further questions because health care is delivered on a provincial level and, as such, can vary in terms of policies from province to province.

The goal of this proposed study is to evaluate the impact of burnout on physicians, residents, and medical students in Alberta and identify the individual, occupational, and organizational variables that influence burnout dimensions. The results of this study will be used to propose specific interventions and preventative courses of action to mitigate burnout and subsequent impairment among physicians, residents, and medical students in Alberta. Given this overall goal of our study, our specific research questions include:

- What are the prevalence rates and correlates of burnout among medical students, residents, and physicians in Alberta?
- What are the perceptions and experience of respondents about the consequences of burnout on the personal and professional life of medical students, residents, and physicians in Alberta?
- What interventions can be implemented to mitigate burnout and promote health and wellness among medical students, residents and physicians in Alberta?

It is hypothesized that the findings will confirm that the prevalence rates of burnout among trainees and physicians are consistent with those reported in other jurisdictions. Our study will also identify whether personal and organization-related factors contribute to burnout among respondents, and lastly, respondents will identify multiple interventions that can be implemented to promote health and wellness among medical students, residents, and physicians in Alberta.

Methods

Study Setting

This study will be conducted in Alberta, a western province of Canada with a population of 4,286,134 in 2017 [45]. Alberta is divided into 5 administrative health regions, namely the Edmonton, Calgary, Southern, Northern, and Central zones. Health care is administered mainly through Alberta Health Services and Primary Care Networks and Medicentre Clinics. As of December 2018, 10,674 physicians were listed on the in-province registers held by the College of Physicians and Surgeons of Alberta [46]. There are two medical schools in the province that train both medical students and residents: University of Alberta, Faculty of Medicine and Dentistry and University of Calgary, Cumming School of Medicine. As of December 2018, there were 1594 postgraduate medical residents and 1148 medical students registered with the College of Physicians and Surgeons of Alberta [46].

Study Design, Sample Size, and Institutional Review Board Approval

Quantitative and qualitative data collected through self-administered, anonymous, online questionnaires will be used in this cross-sectional provincial study design. The study will be comprised of 3 arms including medical students, residents, and practicing physicians within Alberta. Each arm will have a specific survey oriented towards data collection that is relevant to the study population in question.

For the medical student arm of the study, given the total medical student population in Alberta of 1148 [46], an anticipated sample size of 777 was determined based upon a 95% confidence level

and a margin of error of 2% for prevalence rate estimates for medical student burnout. Similarly, for the resident arm of the study, given the total resident population in Alberta of 1594 [46], an anticipated sample size of 959 was determined based upon a 95% confidence level and a margin of error of 2% for prevalence rate estimates for resident burnout. Finally, for the physician arm of the study, given the total physician population in Alberta of 10,674 [46], an anticipated sample size of 1961 was determined based upon a 95% confidence level and a margin of error of 2% for prevalence rate estimates for physician burnout.

This study will be carried out in accordance with the recommendations of the Health Ethics Research Board of the University of Alberta and the Conjoint Health Research Ethics Board of the University of Calgary. The study will also be conducted in accordance with the Declaration of Helsinki (Hong Kong Amendment) and Good Clinical Practice (Canadian Guidelines). All participants will be provided with an online information leaflet and will provide informed consent prior to participation, in accordance with the Declaration of Helsinki. The protocol has received ethical approval by the Health Ethics Research Board of the University of Alberta (reference number Pro00091436) and the Conjoint Health Research Ethics Board of the University of Calgary (REB19-1457).

Data Collection Tools

Data collection tools for each arm of the study were developed based on published literature and questions from previously validated instruments. The general constructs of interest included relevant demographic information, current practice and career planning, general health status, mental health status, and rates of burnout, as well as factors contributing to both burnout and resilience among respondents. The qualitative portion of the study will be gathered from conceptualized themes arising from the results within each of the constructs of interest and its association to burnout. In addition, each survey will include open-ended questions to facilitate qualitative data collection. Standardized measures from which questions were selected and included in the survey were the Maslach Burnout Inventory, Patient Health Questionnaire 9, Canadian Medical Association National Physician Health Survey, Mini Z burnout survey, Professional Fulfillment Index, and two-item Connor-Davidson Resilience Scale [1,44,47-51]. These standardized measures will provide information for the quantitative portion of the study results.

Prior to the finalization of the study survey, the research team sought feedback regarding the developed data collection tools and selected standardized instruments from stakeholders in Alberta, including representatives of the College of Physicians and Surgeons of Alberta's Physician Wellness Program, Alberta Medical Association's Physician and Family Support Program, University of Alberta Faculty of Medicine & Dentistry's Offices of Advocacy & Wellbeing, and University of Calgary Department of Medicine's Physician Wellness and Vitality Program. Feedback from these stakeholders was used to revise the data collection tools to be used in the study. The revised baseline data collection tools were then pretested on 5 randomly selected representatives from each arm of the study. Feedback from the pretest was further used to revise the data collection tool before it was finally adopted for use in the study.

Eligibility, Data Collection Procedures, and Analysis

All medical students and residents at the University of Alberta and the University of Calgary as well as physicians registered with the College of Physicians and Surgeons of Alberta and practicing medicine in the province of Alberta are eligible to participate in the study. Data collection will occur online using "Survey Select," an electronic survey platform, which is hosted by Alberta Health Services (the provincial health authority).

A link to the online information leaflet, consent page, and survey questionnaires will be emailed to all medical students and residents in Alberta through the mailing list of the Offices of the Undergraduate and Post Graduate Medical Education at the University of Alberta and University of Calgary. In order to reach physician respondents, the link will be sent through multiple sources, including the Alberta Medical Association, Alberta Health Services, Primary Care Networks, and Medicentre organizations. Data collection will occur over 4 weeks, and reminder emails will be sent to all eligible respondents weekly. Quantitative data will be analyzed using SPSS 20.0 (IBM Corp, Armonk, NY). Descriptive and inferential statistics will be used to describe demographic characteristics and study variables. Qualitative data will be analyzed using manual thematic analysis.

Results

We anticipate that the study recruitment will begin in September 2020, with 4 weeks of data collection, and study findings will be available within 12 months following completion of data collection. It is expected that the findings will confirm that the prevalence rates of burnout among trainees and physicians are consistent with those reported in other jurisdictions. In addition, we expect to uncover specific contributors to burnout, which will serve as an opportunity for identification of meaningful solutions from the respondents' perspectives.

We intend to disseminate the research findings at several levels, including trainees, physicians, academics, researchers, and health care organizations, as well as membership associations and licensing colleges. This information will be disseminated to academics and stakeholders through research forums and peer-reviewed journals. The expected findings will become available to trainees and physicians through the same communication channels used to provide the initial link for data collection purposes. Namely, this will include the Offices of the Undergraduate and Postgraduate Medical Education at the University of Alberta and University of Calgary, Alberta Medical Association, Alberta Health Services, Primary Care Networks, and Medicentre organizations.

Discussion

This study will contribute to and build on current knowledge by identifying rates of burnout among stages of training and practice in the medical profession, uncovering specific contributors to burnout, and identifying potentially meaningful solutions from the respondents' perspectives. This study is

relatively unique for the following reasons: physicians at multiple stages of training and practice are included; both antecedents to burnout and possible interventions will be examined; and the study population will be broad but within the context of the local health system in Alberta.

It is estimated that 30% of medical trainees and physicians in Alberta will experience burnout based on national study data [44]. Poor work-life balance in the setting of increasingly challenging work environments and additional psychosocial stressors outside of medicine can make physicians and trainees particularly vulnerable to burnout [12,24,38,39]. Both work-related factors (eg, area of practice, hours worked per week, lower autonomy, and number of call shifts per week) and non-work related factors (eg, age, gender, number of dependents, and fewer years of experience) have been shown to affect burnout rates [7,40,41].

While the personal impact of burnout on physicians and medical trainees is substantial, the subsequent costs to patients and the health care system in general are most concerning. Trainees and physicians experiencing burnout are more likely to tolerate unprofessional behavior and encounter negative patient-physician relationships as well as negative interactions with other health care professionals, thereby considering leaving the medical profession mid-career [14,20-22]. Such erosion in practitioners' confidence can subsequently place additional burdens on the health care system, with human resource shortages and expanded waiting times [26,52]. A Canadian study published in 2014 estimated that the total cost of burnout from physicians retiring prematurely or reducing clinical hours was approximately Can $213.1 million, revealing the financial extent and impact of the sequelae of burnout on the health system [26].

Unfortunately, there is a paucity of data regarding interventions that have proven beneficial in impacting burnout rates past the intervention period. Mentorship programs and good occupational leadership have been proposed to potentially influence the wellbeing and satisfaction of individual physicians working in health care organizations [53]. Nevertheless, interventions aimed at reducing burnout will need to be tailored to the stage of training and practice, which will require recognition of the factors contributing to burnout at each stage.

In summary, although physician health and wellness are steadily gaining recognition as a serious issue, gaps remain regarding how burnout affects trainees and physicians in Alberta, the consequences that derive from burnout, and the interventions used to reduce burnout. What appears to be needed is an understanding of the factors contributing to burnout in the context of local health care systems. Valuable insight can be gained from the perceptions and experiences of trainees and physicians in Alberta. These can in turn inform organizational strategies to mitigate burnout and promote wellness within the medical profession, thereby impacting patient care and reducing the burden on health care systems.

Acknowledgments

This work is being funded by the Psychiatry Residents Research Fund held at the Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta.

Conflicts of Interest

None declared.

References

1. Maslach C, Jackson S, Leiter M. Maslach Burnout Inventory Manual. 3rd edition. Palo Alto, CA: Consulting Psychologists Press; 1996:9789994892242.
2. Rotenstein LS, Torre M, Ramos MA, Rosales RC, Guille C, Sen S, et al. Prevalence of Burnout Among Physicians: A Systematic Review. JAMA 2018 Sep 18;320(11):1131-1150 [FREE Full text] [doi: 10.1001/jama.2018.12777] [Medline: 30326495]
3. Cordes CL, Dougherty TW. A Review and an Integration of Research on Job Burnout. Acad Manage Rev 1993 Oct;18(4):621-656. [doi: 10.5465/amr.1993.9402210153]
4. Maslach C. Job Burnout: New Directions in Research and Intervention. Curr Dir Psychol Sci 2003 Oct;12(5):189-192. [doi: 10.1111/1467-8721.01258]
5. Maslach C, Jackson SE. The Measurement of Experienced Burnout. J Organ Behav 1981 Apr;2(2):99-113. [doi: 10.1002/job.4030020205]
6. Maslach C, Leiter MP. Early Predictors of Job Burnout and Engagement. J Appl Psychol 2008 May;93(3):498-512. [doi: 10.1037/0021-9010.93.3.498] [Medline: 18457483]
7. Romani M, Ashkar K. Burnout Among Physicians. Libyan J Med 2014 Feb;9:23556 [FREE Full text] [doi: 10.3402/ljm.v9.23556] [Medline: 24560380]
8. Norlund S, Reuterwall C, Höög J, Lindahl B, Janlert U, Birgander LS. Burnout, Working Conditions and Gender - Results from the Northern Sweden MONICA Study. BMC Public Health 2010 Jun 09;10:326 [FREE Full text] [doi: 10.1186/1471-2458-10-326] [Medline: 20534136]

9. Ahola K, Kivimäki M, Honkonen T, Virtanen M, Koskinen S, Vahtera J, et al. Occupational Burnout and Medically Certified Sickness Absence: A Population-based Study of Finnish Employees. J Psychosom Res 2008 Feb;64(2):185-193. [doi: 10.1016/j.jpsychores.2007.06.022] [Medline: 18222132]
10. Kant IJ, Bültmann U, Schröer KAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. An Epidemiological Approach to Study Fatigue in the Working Population: The Maastricht Cohort Study. Occup Environ Med 2003 Jun;60 Suppl 1:i32-i39 [FREE Full text] [doi: 10.1136/oem.60.suppl_1.i32] [Medline: 12782745]
11. Hallsten L, Bellaagh K, Gustafsson K. Utbranning i Sverige - en populationsstudie (Burning-out in Sweden - a population study). Stockholm, Sweden: Arbetslivsinstitutet; 2002:9170456399.
12. Grinspoon P. Harvard Health Blog: Harvard Health Publishing. 2018 Jun 22. Physician Burnout Can Affect Your Health URL: https://www.health.harvard.edu/blog/physician-burnout-can-affect-your-health-2018062214093 [accessed 2020-01-13]
13. Dyrbye LN, West CP, Satele D, Boone S, Tan L, Sloan J, et al. Burnout Among U.S. Medical Students, Residents, and Early Career Physicians Relative to the General U.S. Population. Acad Med 2014 Mar;89(3):443-451 [FREE Full text] [doi: 10.1097/ACM.0000000000000134] [Medline: 24448053]
14. Dyrbye LN, Massie FS, Eacker A, Harper W, Power D, Durning SJ, et al. Relationship Between Burnout and Professional Conduct and Attitudes Among US Medical Students. JAMA 2010 Sep 15;304(11):1173-1180. [doi: 10.1001/jama.2010.1318] [Medline: 20841530]
15. Paro HBMS, Silveira PSP, Perotta B, Gannam S, Enns SC, Giaxa RRB, et al. Empathy Among Medical Students: Is There a Relation with Quality of Life and Burnout? PLoS One 2014;9(4):e94133 [FREE Full text] [doi: 10.1371/journal.pone.0094133] [Medline: 24705887]
16. Thomas NK. Resident Burnout. JAMA 2004 Dec 15;292(23):2880-2889. [doi: 10.1001/jama.292.23.2880] [Medline: 15598920]
17. Brazeau CMLR, Schroeder R, Rovi S, Boyd L. Relationships Between Medical Student Burnout, Empathy, and Professionalism Climate. Acad Med 2010 Oct;85(10 Suppl):S33-S36. [doi: 10.1097/ACM.0b013e3181ed4c47] [Medline: 20881699]
18. Jackson ER, Shanafelt TD, Hasan O, Satele DV, Dyrbye LN. Burnout and Alcohol Abuse/Dependence Among U.S. Medical Students. Acad Med 2016 Sep;91(9):1251-1256. [doi: 10.1097/ACM.0000000000001138] [Medline: 26934693]
19. Dyrbye LN, Thomas MR, Massie FS, Power DV, Eacker A, Harper W, et al. Burnout and Suicidal Ideation Among U.S. Medical Students. Ann Intern Med 2008 Sep 02;149(5):334-341. [doi: 10.7326/0003-4819-149-5-200809020-00008] [Medline: 18765703]
20. West CP, Huschka MM, Novotny PJ, Sloan JA, Kolars JC, Habermann TM, et al. Association of Perceived Medical Errors with Resident Distress and Empathy: A Prospective Longitudinal Study. JAMA 2006 Sep 06;296(9):1071-1078. [doi: 10.1001/jama.296.9.1071] [Medline: 16954486]
21. West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of Resident Fatigue and Distress with Perceived Medical Errors. JAMA 2009 Sep 23;302(12):1294-1300. [doi: 10.1001/jama.2009.1389] [Medline: 19773564]
22. West CP, Shanafelt TD, Kolars JC. Quality of Life, Burnout, Educational Debt, and Medical Knowledge Among Internal Medicine Residents. JAMA 2011 Sep 07;306(9):952-960. [doi: 10.1001/jama.2011.1247] [Medline: 21900135]
23. van der Heijden F, Dillingh G, Bakker A, Prins J. Suicidal Thoughts Among Medical Residents with Burnout. Arch Suicide Res 2008;12(4):344-346. [doi: 10.1080/13811110802325349] [Medline: 18828037]
24. Campbell J, Prochazka AV, Yamashita T, Gopal R. Predictors of Persistent Burnout in Internal Medicine Residents: A Prospective Cohort Study. Acad Med 2010 Oct;85(10):1630-1634. [doi: 10.1097/ACM.0b013e3181f0c4e7] [Medline: 20881685]
25. Shanafelt TD, Hasan O, Dyrbye LN, Sinsky C, Satele D, Sloan J, et al. Changes in Burnout and Satisfaction With Work-Life Balance in Physicians and the General US Working Population Between 2011 and 2014. Mayo Clin Proc 2015 Dec;90(12):1600-1613. [doi: 10.1016/j.mayocp.2015.08.023] [Medline: 26653297]
26. Dewa CS, Jacobs P, Thanh NX, Loong D. An Estimate of the Cost of Burnout on Early Retirement and Reduction in Clinical Hours of Practicing Physicians in Canada. BMC Health Serv Res 2014 Jun 13;14:254 [FREE Full text] [doi: 10.1186/1472-6963-14-254] [Medline: 24927847]
27. Oreskovich MR, Kaups KL, Balch CM, Hanks JB, Satele D, Sloan J, et al. Prevalence of Alcohol Use Disorders Among American Surgeons. Arch Surg 2012 Feb;147(2):168-174. [doi: 10.1001/archsurg.2011.1481] [Medline: 22351913]
28. Oreskovich MR, Shanafelt T, Dyrbye LN, Tan L, Sotile W, Satele D, et al. The Prevalence of Substance Use Disorders in American Physicians. Am J Addict 2015 Jan;24(1):30-38. [doi: 10.1111/ajad.12173] [Medline: 25823633]
29. Anagnostopoulos F, Demerouti E, Sykioti P, Niakas D, Zis P. Factors Associated with Mental Health Status of Medical Residents: A Model-guided Study. J Clin Psychol Med Settings 2015 Mar;22(1):90-109. [doi: 10.1007/s10880-014-9415-2] [Medline: 25554496]
30. Dahlin M, Joneborg N, Runeson B. Performance-based Self-esteem and Burnout in a Cross-sectional Study of Medical Students. Med Teach 2007 Feb;29(1):43-48. [doi: 10.1080/01421590601175309] [Medline: 17538833]
31. Kassam A, Horton J, Shoimer I, Patten S. Predictors of Well-Being in Resident Physicians: A Descriptive and Psychometric Study. J Grad Med Educ 2015 Mar;7(1):70-74 [FREE Full text] [doi: 10.4300/JGME-D-14-00022.1] [Medline: 26217426]

32. Klein J, Grosse Frie K, Blum K, von dem Knesebeck O. Burnout and Perceived Quality of Care Among German Clinicians in Surgery. Int J Qual Health Care 2010 Dec;22(6):525-530. [doi: 10.1093/intqhc/mzq056] [Medline: 20935011]
33. Schaufeli W, Van Dierendonck D. Utrechtse Burnout Schaal (UBOS): Handleiding. Utrecht Burnout Scale, Manual. Lisse, The Netherlands: Swets & Zeitlinger; 2000.
34. Demerouti E, Bakker A. Chapter 6. The Oldenburg Burnout Inventory: A Good Alternative to Measure Burnout and engagement. Handbook of Stress and Burnout in Health Care. New York, NY: Nova Science Publishers; 2008:9781604565003.
35. Kristensen TS, Borritz M, Villadsen E, Christensen KB. The Copenhagen Burnout Inventory: A New Tool for the Assessment of Burnout. Work & Stress 2005 Jul;19(3):192-207. [doi: 10.1080/02678370500297720]
36. Swider BW, Zimmerman RD. Born to Burnout: A Meta-analytic Path Model of Personality, Job Burnout, and Work Outcomes. J Vocat Behav 2010 Jun;76(3):487-506. [doi: 10.1016/j.jvb.2010.01.003]
37. McManus IC, Keeling A, Paice E. Stress, Burnout and Doctors' Attitudes to Work are Determined by Personality and Learning Style: A Twelve Year Longitudinal Study of UK Medical Graduates. BMC Med 2004 Aug 18;2:29 [FREE Full text] [doi: 10.1186/1741-7015-2-29] [Medline: 15317650]
38. Dyrbye LN, Thomas MR, Huntington JL, Lawson KL, Novotny PJ, Sloan JA, et al. Personal Life Events and Medical Student Burnout: A Multicenter Study. Acad Med 2006 Apr;81(4):374-384. [Medline: 16565189]
39. Ripp J, Babyatsky M, Fallar R, Bazari H, Bellini L, Kapadia C, et al. The Incidence and Predictors of Job Burnout in First-year Internal Medicine Residents: A Five-institution Study. Acad Med 2011 Oct;86(10):1304-1310. [doi: 10.1097/ACM.0b013e31822c1236] [Medline: 21869661]
40. Adam S, Mohos A, Kalabay L, Torzsa P. Potential Correlates of Burnout Among General Practitioners and Residents in Hungary: The Significant Role of Gender, Age, Dependant Care and Experience. BMC Fam Pract 2018 Dec 12;19(1):193 [FREE Full text] [doi: 10.1186/s12875-018-0886-3] [Medline: 30541461]
41. Shanafelt TD, Balch CM, Bechamps GJ, Russell T, Dyrbye L, Satele D, et al. Burnout and Career Satisfaction Among American Surgeons. Ann Surg 2009 Sep;250(3):463-471. [doi: 10.1097/SLA.0b013e3181ac4dfd] [Medline: 19730177]
42. Canadian MA. CMA PolicyBase: Ottawa, Canada. 2017. Background to CMA Policy on Physician Health URL: https://policybase.cma.ca/en/viewer?file=%2fdocuments%2fPolicypdf%2fPD18-01S.pdf#phrase=false [accessed 2020-01-13]
43. Canadian MA. CMA PolicyBase: Ottawa, Canada. 2017. CMA Policy on Physician Health URL: https://policybase.cma.ca/en/viewer?file=%2fdocuments%2fPolicypdf%2fPD18-01.pdf#phrase=false [accessed 2020-01-13]
44. Canadian MA. CMA National Physician Health Survey: A National Snapshot October. 2018. URL: https://www.cma.ca/sites/default/files/2018-11/nph-survey-e.pdf [accessed 2020-01-13]
45. Statistics C. Canada at a Glance. Minister of Industry: Statistics Canada; 2018 Mar 27. CANSIM table 051-0005 URL: https://www150.statcan.gc.ca/n1/en/pub/12-581-x/12-581-x2018000-eng.pdf?st=BAoZQwhW [accessed 2020-01-13]
46. CPSA. Quarterly Report Q4.: College of Physicians and Surgeons of Alberta; 2018. Quarterly Update Physician Resources 2018 Oct 1 to 2018 Dec 31 URL: http://www.cpsa.ca/wp-content/uploads/2019/01/Quarterly-Report-Q4-2018.pdf [accessed 2020-01-13]
47. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a Brief Depression Severity Measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
48. Linzer M, Konrad TR, Douglas J, McMurray JE, Pathman DE, Williams ES, et al. Managed Care, Time Pressure, and Physician Job Satisfaction: Results from the Physician Worklife Study. J Gen Intern Med 2000 Jul;15(7):441-450 [FREE Full text] [Medline: 10940129]
49. Linzer M, Poplau S, Grossman E, Varkey A, Yale S, Williams E, et al. A Cluster Randomized Trial of Interventions to Improve Work Conditions and Clinician Burnout in Primary Care: Results from the Healthy Work Place (HWP) Study. J Gen Intern Med 2015 Aug;30(8):1105-1111 [FREE Full text] [doi: 10.1007/s11606-015-3235-4] [Medline: 25724571]
50. Trockel M, Bohman B, Lesure E, Hamidi MS, Welle D, Roberts L, et al. A Brief Instrument to Assess Both Burnout and Professional Fulfillment in Physicians: Reliability and Validity, Including Correlation with Self-Reported Medical Errors, in a Sample of Resident and Practicing Physicians. Acad Psychiatry 2018 Feb;42(1):11-24 [FREE Full text] [doi: 10.1007/s40596-017-0849-3] [Medline: 29196982]
51. Vaishnavi S, Connor K, Davidson JRT. An Abbreviated Version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: Psychometric Properties and Applications in Psychopharmacological Trials. Psychiatry Res 2007 Aug 30;152(2-3):293-297 [FREE Full text] [doi: 10.1016/j.psychres.2007.01.006] [Medline: 17459488]
52. Dewa CS, Loong D, Bonato S, Trojanowski L. The Relationship Between Physician Burnout and Quality of Healthcare in Terms of Safety and Acceptability: A Systematic Review. BMJ Open 2017 Jun 21;7(6):e015141 [FREE Full text] [doi: 10.1136/bmjopen-2016-015141] [Medline: 28637730]
53. Shanafelt TD, Gorringe G, Menaker R, Storz KA, Reeves D, Buskirk SJ, et al. Impact of Organizational Leadership on Physician Burnout and Satisfaction. Mayo Clin Proc 2015 Apr;90(4):432-440. [doi: 10.1016/j.mayocp.2015.01.012] [Medline: 25796117]

Edited by G Eysenbach; submitted 17.09.19; peer-reviewed by D Gokcay, E Kleiman; comments to author 02.12.19; revised version received 26.01.20; accepted 21.02.20; published 17.04.20.

Please cite as:
Kim E, Mallett R, Hrabok M, Yang YA, Moreau C, Nwachukwu I, Kravtsenyuk M, Abba-Aji A, Li D, Agyapong VIO
Reducing Burnout and Promoting Health and Wellness Among Medical Students, Residents, and Physicians in Alberta: Protocol for a Cross-Sectional Questionnaire Study
JMIR Res Protoc 2020;9(4):e16285
URL: http://www.researchprotocols.org/2020/4/e16285/
doi:10.2196/16285
PMID:32301742

Protocol

Use of Smart Technology for the Early Diagnosis of Complications After Cardiac Surgery: The Box 2.0 Study Protocol

Tom E Biersteker[1], MD; Mark J Boogers[1], MD, PhD; Robert AF de Lind van Wijngaarden[2], MD, PhD; Rolf HH Groenwold[3], MD, PhD; Serge A Trines[1], MD, PhD; Anouk P van Alem[4], MD, PhD; Charles JHJ Kirchhof[5], MD, PhD; Nicolette van Hof[1], MANP; Robert JM Klautz[2], MD, PhD; Martin J Schalij[1], MD, PhD; Roderick W Treskes[1], PhD

[1]Department of Cardiology, Leiden University Medical Center, Leiden, Netherlands

[2]Department of Cardiothoracic Surgery, Leiden University Medical Center, Leiden, Netherlands

[3]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands

[4]Department of Cardiology, Haaglanden Medisch Centrum, Den Haag, Netherlands

[5]Department of Cardiology, Alrijne Ziekenhuis, Leiderdorp, Netherlands

Corresponding Author:
Tom E Biersteker, MD
Department of Cardiology
Leiden University Medical Center
Albinusdreef 2
Leiden, 2333 ZA
Netherlands
Phone: 31 715262020
Email: t.e.biersteker@lumc.nl

Abstract

Background: Atrial fibrillation (AF), sternal wound infection, and cardiac decompensation are complications that can occur after cardiac surgery. Early detection of these complications is clinically relevant, as early treatment is associated with better clinical outcomes. Remote monitoring with the use of a smartphone (mobile health [mHealth]) might improve the early detection of complications after cardiac surgery.

Objective: The primary aim of this study is to compare the detection rate of AF diagnosed with an mHealth solution to the detection rate of AF diagnosed with standard care. Secondary objectives include detection of sternal wound infection and cardiac decompensation, as well as assessment of quality of life, patient satisfaction, and cost-effectiveness.

Methods: *The Box 2.0* is a study with a prospective intervention group and a historical control group for comparison. Patients undergoing cardiac surgery at Leiden University Medical Center are eligible for enrollment. In this study, 365 historical patients will be used as controls and 365 other participants will be asked to receive either *The Box 2.0* intervention consisting of seven home measurement devices along with a video consultation 2 weeks after discharge or standard cardiac care for 3 months. Patient information will be analyzed according to the intention-to-treat principle. *The Box 2.0* devices include a blood pressure monitor, thermometer, weight scale, step count watch, single-lead electrocardiogram (ECG) device, 12-lead ECG device, and pulse oximeter.

Results: The study started in November 2018. The primary outcome of this study is the detection rate of AF in both groups. Quality of life is measured with the five-level EuroQol five-dimension (EQ-5D-5L) questionnaire. Cost-effectiveness is calculated from a society perspective using prices from Dutch costing guidelines and quality of life data from the study. In the historical cohort, 93.9% (336/358) completed the EQ-5D-5L and patient satisfaction questionnaires 3 months after cardiac surgery.

Conclusions: The rationale and design of a study to investigate mHealth devices in postoperative cardiac surgery patients are presented. The first results are expected in September 2020.

Trial Registration: ClinicalTrials.gov NCT03690492; http://clinicaltrials.gov/show/NCT03690492

International Registered Report Identifier (IRRID): DERR1-10.2196/16326

(JMIR Res Protoc 2020;9(4):e16326) doi:10.2196/16326

KEYWORDS

mHealth; cardiac surgery; atrial fibrillation; postoperative care; ambulatory electrocardiography

Introduction

Over the past decades, major advances have been made to improve the safety of cardiac surgery in order to decrease the risk of adverse events [1]. However, a series of life-threatening events can occur after cardiac surgery. Even patients who are discharged in a clinically stable condition are still at risk for the development of complications at home. The most frequently occurring postoperative complications are cardiac decompensation, late tamponade, and rhythm disturbances, such as atrial fibrillation (AF). In approximately 25% of all patients, one or more of these complications occur, and they are mostly noted in the ward at day 2 or 3 after surgery [2-4]. At day 7 or later, postoperative atrial fibrillation (POAF) is diagnosed in approximately 15% of all cases [5]. Sternal wound infection occurs in 3% to 5% of cases, with mediastinitis occurring in 1% to 2% of all cases [6-8]. Late-onset sternal wound infection is defined as sternal wound infection occurring 14 days or later after the initial surgery, and it is relatively as frequent as early-onset sternal wound infection [9].

Early detection of these complications is of vital importance as there can be various issues (eg, untreated AF is associated with an increased risk of transient ischemic attack and ischemic stroke) [10,11]. Moreover, with early diagnosis of cardiac decompensation or wound infection, hospital readmission may be prevented [12]. Previous research has shown that patients do not always recognize complications at home, which may delay the diagnosis and thus cause more risk and harm to the patient [13]. Telephone calls as a follow-up strategy have been found to be helpful, but ambulatory diagnosis of postoperative complications remains challenging [14].

Advances in information and communication technologies have led to more possibilities for monitoring patients remotely (telemonitoring) and using a smartphone to provide medical care (mobile health [mHealth]) [15]. For example, smartphone-compatible detectors for cardiovascular disease parameters, such as Kardia (AliveCor Inc, San Francisco, California, USA) and Withings Blood Pressure Monitor (Withings, Issy les Moulineaux, France), have been released in the consumer market [16-19]. These devices are capable of measuring blood pressure (BP), blood oxygen saturation, weight, temperature, and number of daily steps taken, as well as electrocardiogram (ECG) registration, providing patients with the possibility to monitor their vital parameters at home.

In addition to improved monitoring, mHealth may have more benefits. Video conferences between doctors and patients may save time and money for patients, especially in rural areas [20,21]. Unpublished results of *The Box*, an mHealth study in patients after myocardial infarction, which has been registered under clinical trial number NCT02976376 at ClinicalTrials.gov, suggest that remote follow-up is cost-effective [22].

It is hypothesized that smart technology can help with the early diagnosis of late complications and thereby improve quality of care and patient satisfaction after cardiac surgery. The aim of this study is to investigate the clinical effectiveness of a smart technology intervention in patients after cardiac surgery. The rationale and design of this study are presented.

Methods

Study Design

The Box 2.0 is a study with a prospective intervention group and a historical control group for comparison. The study is being conducted at the Department of Cardiothoracic Surgery and Department of Cardiology in Leiden University Medical Center (LUMC), a tertiary care hospital in Leiden, The Netherlands. The study is registered at ClinicalTrials.gov under trial number NCT03690492 and registered at Toetsingonline.nl under number NL65959.058.18, and has been approved by the Medical Ethics Committee of LUMC (P18.110). All procedures are conducted in accordance with the principles of the Declaration of Helsinki (version 10, October 2013) and in accordance with the Dutch Medical Research Involving Human Subjects Act (Wet Medisch-wetenschappelijk Onderzoek met mensen) and Good Clinical Practice. Written offline informed consent will be obtained from all prospective study participants.

Patient Population

Adult patients undergoing coronary artery bypass grafting, valve reconstruction or replacement, surgery of the aortic root or ascending aorta, or other cardiac surgeries performed via median sternotomy, such as atrial or ventricular septal defect closure, a Dor or Morrow procedure, cardiac tumor removal, and surgical treatment of coronary artery anomalies, are eligible for enrolment. Patients with perioperative endocarditis, those with a need for mechanical support with extracorporeal membrane oxygenation (ECMO), an Impella device, or an intra-aortic balloon pump (IABP), and those with an Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS) score of 1 or 2 are excluded. All inclusion and exclusion criteria are listed in Textbox 1.

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria

- Patient undergoes cardiac surgery (coronary artery bypass grafting, valve reconstruction or replacement, aortic root or ascending aortic surgery, or other cardiac surgeries performed via median sternotomy, such as atrial or ventricular septal defect closure, a Dor or Morrow procedure, cardiac tumor removal, and surgical treatment of coronary artery anomalies.
- Patient is able to communicate and is literate in English or Dutch.

Exclusion criteria

- Patient is less than 18 years old.
- Patient is pregnant.
- Patient is considered an incapacitated adult.
- Patient is unwilling to sign the informed consent form.
- Patient undergoes emergency cardiac surgery (Interagency Registry for Mechanically Assisted Circulatory Support score 1 or 2).
- Patient has active endocarditis at the time of operation.
- Patient is on mechanical circulatory support before operation.
- Patient has a ventricular septal rupture.
- Patient undergoes extracorporeal membrane oxygenation or ventricular assist device insertion.

Patient Selection

The historical cohort consists of patients who underwent cardiac surgery at LUMC from January to September 2018. Among the patients, 93.9% (336/358) completed the five-level EuroQol five-dimension (EQ-5D-5L) questionnaire and patient satisfaction questionnaire 3 months after cardiac surgery. These questionnaires are consistent with the questionnaires to be filled out by prospective study patients.

The prospective cohort is selected from November 2018 onwards. Patients eligible for participation are recruited through three different pathways as follows: patients who are electively awaiting cardiac surgery will be approached for participation in the outpatient clinic 4 to 6 weeks before their operation; patients who are admitted to the cardiology ward and awaiting surgery will be approached for participation approximately 3 to 5 days before surgery; and patients from a referring hospital who are transferred to the cardiothoracic surgery ward of LUMC awaiting surgery will be approached for participation approximately 2 days before surgery. If a patient is not included in the study preoperatively, the patient can be asked to participate after the operation until discharge or transfer back to the referring hospital.

Standard Postoperative Care

Fourteen days and 3 months after cardiac surgery, patients are seen by a nurse practitioner (NP), who is supervised by a consultant cardiologist. At 14 days, the sternal wound and, if applicable, the venous graft resection site are checked, and vital parameters, signs of congestion, indications for and side effects of medications, and symptoms of postoperative complications are assessed.

At 3 months, a similar examination is performed during an outpatient clinic visit. In addition, blood samples are taken for measuring kidney function and lipid spectrum. Moreover, a transthoracic echocardiogram is obtained. Patients undergoing a mini-Maze procedure or concomitant radiofrequency ablation will also undergo 24-hour Holter monitoring and physical exercise testing at 3 and 12 months after surgery.

The Box 2.0

The Box 2.0 Intervention

Patients who consent to participate in *The Box 2.0* will receive a box containing a Withings weight scale, Withings BP monitor, Withings activity tracker, Withings thermometer, blood oxygen saturation monitor (Masimo, Irvine, California, USA), and two mobile ECG devices (AliveCor Kardia and CardioSecur; Personal MedSystems, Frankfurt am Main, Germany). The devices are shown in Figure 1.

Figure 1. The devices of *The Box 2.0*.

All are noninvasive, battery-powered, smartphone-compatible devices. They are Conformité Européenne marked and approved by the Food and Drug Administration and can be obtained in retail stores in the European Union and the United States. No assistance is required to install and use the devices; however, professional assistance is provided by a technical LUMC employee when a patient requires help.

A smartphone with iOS (Apple Computers, Cupertino, California, USA) or Android Operating System (Google, Mountain View, California, USA) is required to use the devices, as they communicate with a dedicated mobile app on the smartphone, which can be downloaded from the AppStore (iOS) or PlayStore (Android). Measurement data are stored on the smartphone and uploaded to the app manufacturer's servers (located in Europe for Withings and CardioSecur and the United States for AliveCor and Masimo). An internet connection is required to synchronize the data, although measurements can be obtained when the smartphone is offline. The data are uploaded to the corresponding servers when the smartphone reconnects to the internet.

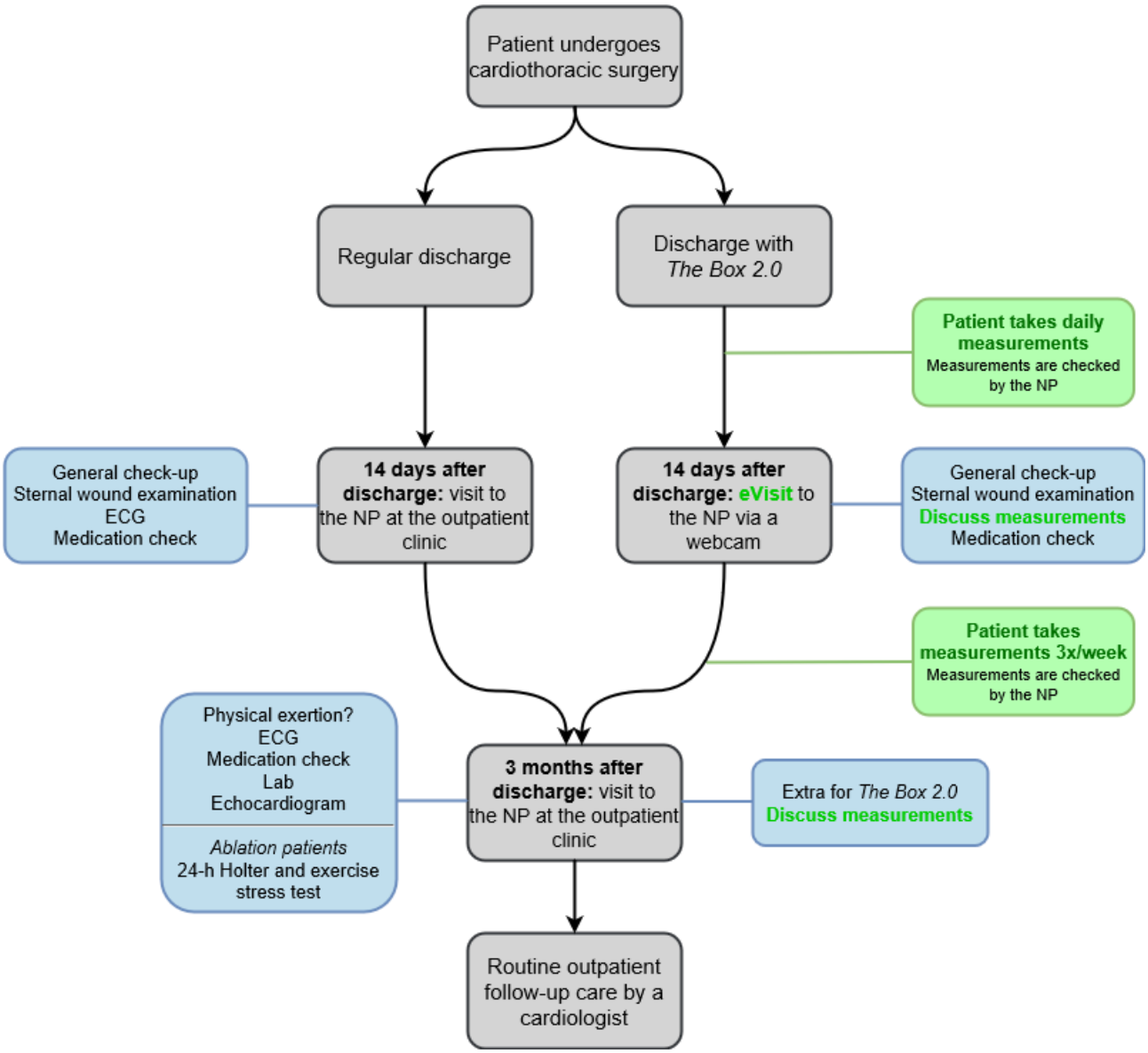
Patients receive *The Box 2.0* before discharge. A technical support employee of LUMC instructs the patients on how to install and use the devices on their mobile phones, including registering necessary accounts. This employee also manages a help desk service that patients can reach for technical issues with their devices. Patients are also provided with written information and an instruction video. Patients who do not own a smartphone with either iOS or Android are provided with a Samsung J3 smartphone (Samsung, Seoul, South Korea). Patients are instructed to use their own internet service (WiFi or mobile network). No mobile data network plan is provided with the smartphone. Additionally, to be able to use the accounts, an email address is required. Patients are provided with a randomly generated email address from the @hlc.nl domain, which is owned and maintained by LUMC.

Patients are instructed to record their BP, weight, temperature, step count, blood oxygen saturation, respiration rate, and single-lead ECG data with AliveCor Kardia on a daily basis for the first 14 days. After 14 days, all measurements except for step count are reduced to thrice a week. Additionally, patients record a CardioSecur ECG weekly. In case of a possible rhythm disturbance, which is either felt by the patient or diagnosed with the single-lead ECG, patients obtain an additional CardioSecur ECG.

Moreover, the first visit to the outpatient clinic is replaced by an electronic visit, in which the patient communicates with the NP via Webcamconsult (Webcamconsult, Bergen op Zoom, The Netherlands), a secured video connection. The same topics as those during a regular visit are discussed, and the sternal wound is checked. Therefore, patients are advised to wear a dress shirt when establishing the video connection, so that no private body parts are shown.

An overview of regular follow-up and follow-up with *The Box 2.0* is shown in Figure 2.

Figure 2. Flowchart for the outpatient clinic follow-up. Left: standard outpatient care without the mHealth intervention. Right: outpatient care with the mHealth intervention. Blue text boxes show the standard topics at 14-day and 3-month consultations. Green text boxes and text highlight the mHealth intervention. ECG: electrocardiogram; NP: nurse practitioner.



Blood Pressure Monitor

The BP monitor (Withings Blood Pressure Monitor) is a validated [23] smartphone-compatible device that allows users to measure systolic and diastolic BP and heart rate. Depending on user preference, the device is placed around the left or right arm, and after pushing a button on the cuff, the device connects to the smartphone via Bluetooth. A measurement takes an average of 20 seconds, and inflation and deflation are automated and initiated via the dedicated Health Mate app for iOS and Android. The results are shown on the smartphone screen and automatically uploaded to Withings servers.

Weight Scale

The weight scale (Withings Body) measures weight, fat percentage, heart rate, and ambient CO_2. By footprint recognition, all measurements are automatically saved under each user's personal account. This way, measurements of users cannot be confused. The results are shown on the weight scale screen and automatically uploaded to Withings servers.

Thermometer

The thermometer (Withings Thermo) is able to measure the temperature of the temporal artery by scanning the forehead. Additionally, a measurement of the dorsal auricular artery can be obtained by placing the device on the soft tissue underneath the ear. The highest measurement will be used. The results are shown on the thermometer and automatically uploaded to Withings servers.

Activity Tracker

The activity tracker (Withings Steel) automatically tracks the number of daily steps taken and the duration and quality of sleep. The device is a small watch and can also be used as such. Results are automatically uploaded to Withings servers.

Blood Oxygen Saturation Monitor

The oxygen saturation monitor (Masimo MightySatRx) tracks blood oxygen saturation, as well as heart rate, respiration rate, plethysmography variability index, and peripheral perfusion index. The device is placed on a finger, and measurements start automatically after starting the app. Masimo claims that its device and algorithm work on well-circulated and less well-circulated fingers. Measurements are automatically saved to a comma-separated file on the smartphone. This file can then be shared by email. No measurement data are saved on Masimo servers.

Single-Lead Electrocardiogram Device

The single-lead ECG device (AliveCor Kardia) uses two electrodes to allow a user to record a 30-second lead-I ECG. The device connects to the AliveCor app via an ultrasound signal, and to record an ECG, the patient places two fingers of both hands on the electrodes while the device is held next to the smartphone at a maximum of 30 cm. The device can also be attached to the back of the smartphone.

The ECG signal is converted to a live single-lead ECG that is shown on the smartphone screen [24]. After the measurement is completed, the AliveCor algorithm calculates the R-R intervals and reports either a normal ECG, possible AF, or undetermined finding [16]. Before saving the ECG, the patient can save notes. The algorithm has a 70% to 97% sensitivity and 98% to 99% specificity for the detection of AF [25-27].

EASI-Derived 12-Lead Electrocardiogram Device

The CardioSecur ECG device uses the five-electrode EASI lead system described by Dower et al [28], removing the ground electrode. It is hypothesized to detect myocardial ischemia and have an amount of baseline noise like that of Mason-Likar ECG [17]. This possibly allows for relatively noise-free rhythm detection. As no external validation studies have been performed, this study will not use the CardioSecur device for the detection of myocardial ischemia.

Patients receive their first CardioSecur ECG in the hospital preoperatively. The four leads are connected to the chest of the patient. One lead is placed on the upper part of the sternum, one lead is placed directly underneath the sternum, and two leads are placed on the mid-axillary line at the same height as that of the lower placed frontal lead. After registration of the first ECG, patients can take their own ECG at home. The ECG is uploaded to CardioSecur servers and sent to the hospital electronic medical record in portable document format (PDF).

Figures 3 and 4 show examples of an ECG obtained with CardioSecur and a single-lead ECG obtained with AliveCor Kardia, respectively.

Figure 3. An electrocardiogram obtained with the CardioSecur device.

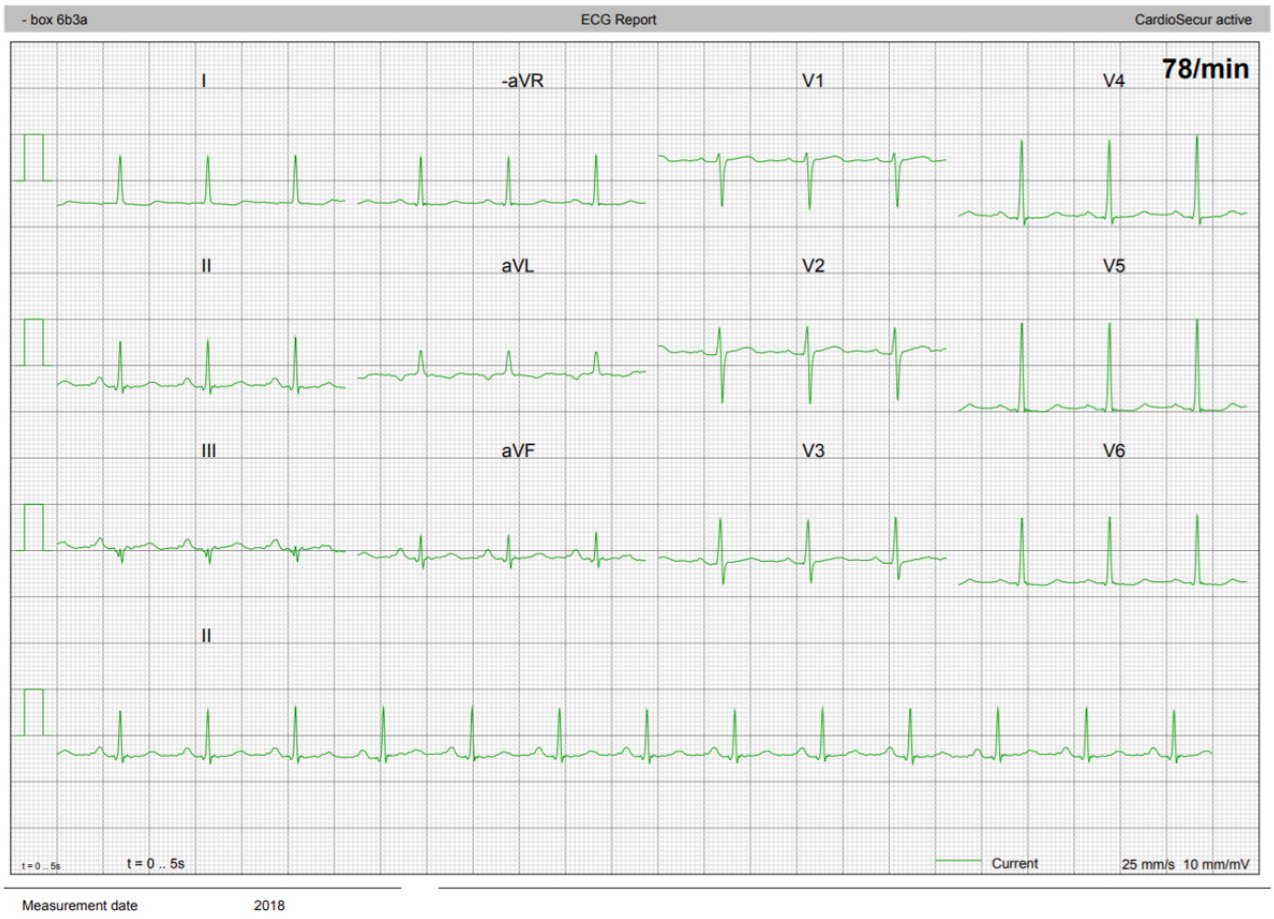
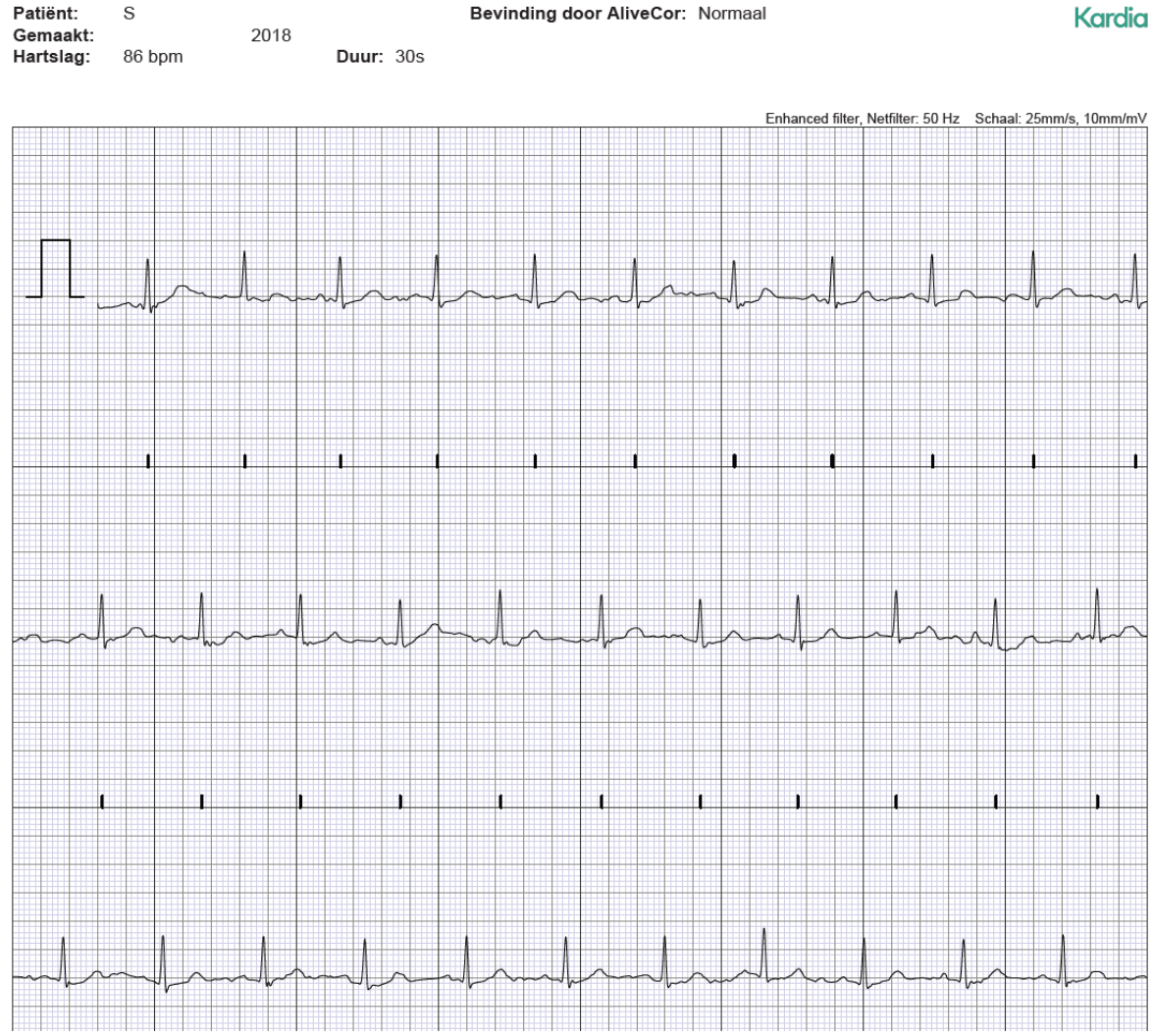


Figure 4. A lead-I electrocardiogram obtained with the AliveCor Kardia device.



Reasons to Contact Patients

The sent-in data are checked daily by a dedicated NP. Data irregularities trigger an automated alarm. When irregularities are noted, patients are contacted within 48 hours after the data have been sent in. The definitions of data irregularities are shown in Multimedia Appendix 1.

Reasons to Adjust the Therapeutic Regimen

According to the sent-in data, the NP can change medication or discuss the data with the patient's treating physician. Detecting an atrial rhythm disturbance with the CardioSecur may thus result in adjusting the therapeutic regiment. This is left to the discretion of the treating physician.

Nonadherence and Study Withdrawal

Patients are required to frequently send in their measurements. If they do not send measurements from their devices for 21 consecutive days, they are considered nonadherent. This is the same policy as in the previous *The Box* study [22]. If a patient is considered nonadherent, a standardized email is sent, urging them to contact the hospital in case of technical difficulties or objections. A specialized technology employee can provide assistance. If, after this phone call, the patient starts sending in measurements, the patient is again considered adherent. If nonadherence reoccurs, a standardized email is sent out. If patients do not start sending in new measurements, they are not approached again and are considered definitively nonadherent. These patients are not excluded from the study and are still followed up according to *The Box 2.0* protocol. If any study patients notify the hospital that they would like to have regular outpatient clinic visits, they are followed up via the regular follow-up protocol. Patients who withdraw from regular follow-up are considered lost to follow-up.

Privacy of Study Participants

To anonymize the data, patients receive a randomly generated email address from the @hlc.nl domain, which is owned and maintained by LUMC. They can use this email address to create their Withings and CardioSecur accounts, as well as send the comma-separated file (created by using the Masimo device) to LUMC. The corresponding names are kept in a separate password-protected database.

Results

Outcomes

The study started recruitment in November 2018. The primary outcome measure of this study in patients undergoing cardiac

surgery is to investigate if POAF is diagnosed more often in patients who are followed up with smart technology as compared with standard care, with measurements until the 3-month outpatient clinic visit. POAF can be diagnosed before or after discharge. As this mHealth study does not impact the AF detection rate before discharge, the primary outcome only involves POAF diagnosed after discharge. In LUMC, the mean duration of hospital stay after cardiac surgery is 8 days. POAF diagnosed before discharge is reported as a baseline characteristic.

AF is defined as an episode of irregular heart rhythm, without detectable P waves, lasting more than 30 seconds [29]. AF can be detected at the outpatient clinic, at the emergency department, or at home via CardioSecur in *TheBox 2.0* patients. As taking an ECG with CardioSecur takes longer than 30 seconds, we defined POAF or atrial flutter as a CardioSecur ECG that shows this rhythm disturbance.

The primary endpoint is evaluated by a clinical event committee that is blinded to patient data. The committee consists of two cardiologists, not otherwise involved in the project, who review the data independently. In case of disagreement, a third cardiologist is involved to reach a decision.

Secondary outcomes of this study include the diagnosis of cardiac decompensation or sternal wound infection. Sternal wound infection is defined by the guidelines of the Center for Disease Control and Prevention, which are most commonly used. For deep sternal wound infection, fever (temperature >38.5°C), sternal instability, or chest discomfort has to be present in combination with either purulent drainage from the sternal wound or mediastinal widening on radiography [30]. Cardiac decompensation is defined as a clinical syndrome in which a structural or functional disorder of the heart impairs the capacity of the ventricle to eject or fill with blood at physiologic filling pressures, therefore limiting the ability of the patient to exercise or carry out activities of daily living without symptoms of dyspnea or fatigue [31]. All outcomes are detailed in Textbox 2 [32,33].

Textbox 2. Primary and secondary outcome measures.

Primary study parameter

- Detection of atrial fibrillation

Secondary study parameters

- Detection of sternal wound infection
- Detection of cardiac decompensation
- Time to detection of atrial fibrillation
- Time to detection of sternal wound infection
- Time to detection of cardiac decompensation
- Quality of life (five-level EuroQol five-dimension)
- Patient satisfaction of care
- Overall mortality
- Major adverse cardiac events: cardiac death, myocardial infarction, cardiac tamponade, ischemic stroke, or transient ischemic attack, defined as a transient episode of neurological dysfunction lasting for less than 1 hour, which is caused by focal brain or retinal ischemia without recent infarction on cerebral imaging [32].
- Readmission to either the cardiology or cardiothoracic surgery ward
- Total number of cardiology-related visits to the emergency department until 3 months after discharge
- Blood pressure control, which is defined as a systolic blood pressure of <140 mmHg and a diastolic blood pressure of <90 mmHg [33]
- Cost-effectiveness

Questionnaires

All patients, both in the intervention and control groups, are asked to fill out a quality of life questionnaire (EQ-5D-5L) [34] and a short patient satisfaction questionnaire. The patients in the intervention group are also asked to fill out a short satisfaction questionnaire concerning the webcam consultation system and the used devices. The EQ-5D-5L questionnaire is used before surgery and 3 months after surgery. The satisfaction questionnaire is used 3 months after surgery.

Sample Size and Statistical Analysis

An AF detection rate of 15% is expected in the historical group and 25% is expected in the intervention group [12,35]. Power calculation was performed using R software (R Foundation for Statistical Computing, Vienna, Austria) [36]. For the calculation, an α of .05 and a power of 0.90 were chosen. This yielded a sample size of 335 patients, which was increased by 9% to 365 patients in both the intervention and control groups to correct for expected loss to follow-up, leading to a total study size of 730 patients.

Patient information is analyzed according to the intention-to-treat principle, and a per protocol subanalysis is carried out. According to the unpublished data of *The Box* study, it is estimated that 33% of eligible patients will not be willing to take part in *The Box 2.0* study or will stop sending in measurements shortly after inclusion. Accordingly, to prevent selection bias, patients who do not wish to take home measurements with the devices of *The Box 2.0* are included in the intervention arm if they consent to the researchers using their data until 3 months after surgery.

With regard to the secondary endpoints, the diagnoses of sternal wound infection and cardiac decompensation are analyzed in the same way as the primary endpoint. The proportion of patients with controlled BP and the proportion of readmitted patients will be compared using a chi-square test. Logistic regression will be performed to correct for potential confounding variables. The scores of questionnaires (EQ-5D-5L and patient satisfaction questionnaires) and health care utilization will be compared using an independent *t* test. As this study is underpowered to detect differences between groups with respect to major adverse cardiac events (MACEs), only descriptive statistics of this outcome will be reported.

Discussion

Patient Selection

For this study, high-risk patients are excluded. The INTERMACS scale was developed to further categorize New York Heart Association Class IV patients [37]. We use this INTERMACS scale as a clinical tool to grade very ill patients, and we define high-risk patients as those with an INTERMACS score of 1 or 2. We exclude patients receiving a ventricular assist device, those with a possible life-threatening condition at the time of their surgery, such as active endocarditis and ventricular septal rupture, and those needing mechanical support with ECMO, IABP, or an Impella device at the time of the operation. This patient group is excluded from participation, as effects on outcomes, such as mortality, MACEs, and rhythm disturbances, would be incomparable between nonrandomized study groups owing to the high mortality and morbidity rates of high-risk patients [38-44]. Patients undergoing any other type of cardiac surgery are included.

Internet Access

As of 2018, 98% of Dutch households have internet access; more than any other European country [45]. Therefore, generalizability of the results of this study might be limited in countries with low internet access percentages. The expected average age of subjects at the time of inclusion is 65 years. In the Netherlands, 84% of people aged 55 to 65 years, 68% of people aged 65 to 75 years, and 40% of people aged 75 years or older have internet access [46]. To prevent selection bias, an intention-to-treat design is chosen, in which patients without internet access are also enrolled in the intervention group.

Comparability and Study Outcomes

To our knowledge, a study validating mHealth in cardiac surgery patients has not been performed. Pilot results from a randomized controlled trial performed by McElroy et al, which involved 27 intervention patients and 416 control patients who underwent cardiac surgery, showed no relevant results regarding hospital readmission rates when using a digital health kit, which included a tablet with a pulse oximeter, heart rate and blood pressure monitor, and weight scale [12]. In this previous study, the proportion of postoperative AF diagnosis did not significantly differ between the groups (29.6% [8/27] in the intervention group and 15.4% [64/416] in the control group; *P*=.90). This may be due to a correction for a difference at baseline, where having a history of AF differed between the groups (22.2% [6/27] in the intervention group and 11.3% [47/416] in the control group; *P*=.20) [12].

Currently available consumer mHealth devices may make ambulatory diagnosis of rhythm disturbances more accessible. As POAF does not always cause symptoms and thus could remain subclinical, we hypothesize that using these devices in patients after cardiac surgery will greatly increase the diagnosis of POAF.

Owing to the accessible nature of consumer mHealth devices, it is hypothesized that other complications that may arise after cardiac surgery, such as sternal wound infection and cardiac decompensation, will be diagnosed earlier as compared with conventional follow-up. Early detection of these abnormalities could lead to a diminished disease burden among patients and result in more outpatient treatment or shorter clinical treatment, causing increased cost-effectiveness. Moreover, as patients can read their own measurements and follow trends, they may feel more empowered with the use of the devices adopted in this study and may reach a higher quality of life.

Definition of Atrial Rhythm Disturbances in mHealth

Defining POAF detected with mHealth devices is not possible according to current European or American guidelines, as mHealth is currently not a tool for diagnosis. Recent studies have found that, when using AF detection devices, prolonged episodes of AF, atrial flutter, or atrial tachycardia (lasting over 5 minutes in the MOST study and over 6 minutes in the ASSERT study) are independently associated with an increased risk of stroke [47,48]. However, short episodes of up to 20 seconds are not associated with an increased risk [49]. Devices of *The Box 2.0* do not measure heart rhythm in a continuous fashion, and thus, measurements should be compared with spot measurements involving a normal clinical ECG. We have therefore initiated a regional focus group with five cardiologists specialized in rhythm disturbances and two cardiac surgeons. This group defined POAF as AF shown on a CardioSecur ECG during the full 10 seconds when the heart rhythm is registered, which is diagnosed within 3 months after cardiac surgery. Other atrial rhythm disturbances, such as atrial flutter and atrial tachycardia, are addressed in the same fashion. For the treatment of diagnosed rhythm disturbances, the current European guidelines are followed [50].

Conclusion

In summary, a study to investigate mHealth devices in cardiac surgery patients is presented. The first results are expected in September 2020.

Acknowledgments

This is an investigator-initiated study. No financial support has been received from manufacturers, as all devices have been purchased for this study. No manufacturer has a role in the study design, data collection, statistical analysis, or manuscript writing. The Department of Cardiology receives unrestricted fellowship grants from Medtronic, Boston Scientific, and Biotronik.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions of data irregularities.
[DOCX File , 24 KB - resprot_v9i4e16326_app1.docx]

References

1. Diodato M, Chedrawy E. Coronary artery bypass graft surgery: the past, present, and future of myocardial revascularisation. Surg Res Pract 2014;2014:726158 [FREE Full text] [doi: 10.1155/2014/726158] [Medline: 25374960]
2. Ball L, Costantino F, Pelosi P. Postoperative complications of patients undergoing cardiac surgery. Current Opinion in Critical Care 2016;22(4):386-392. [doi: 10.1097/mcc.0000000000000319]
3. Siribaddana S. Cardiac dysfunction in the CABG patient. Curr Opin Pharmacol 2012 Apr;12(2):166-171. [doi: 10.1016/j.coph.2012.01.010] [Medline: 22325854]
4. Maisel WH, Rawn JD, Stevenson WG. Atrial fibrillation after cardiac surgery. Ann Intern Med 2001 Dec 18;135(12):1061-1073. [doi: 10.7326/0003-4819-135-12-200112180-00010] [Medline: 11747385]
5. Mathew JP, Fontes ML, Tudor IC, Ramsay J, Duke P, Mazer CD, Investigators of the Ischemia ResearchEducation Foundation, Multicenter Study of Perioperative Ischemia Research Group. A multicenter risk index for atrial fibrillation after cardiac surgery. JAMA 2004 Apr 14;291(14):1720-1729. [doi: 10.1001/jama.291.14.1720] [Medline: 15082699]
6. Lepelletier D, Perron S, Bizouarn P, Caillon J, Drugeon H, Michaud J, et al. Surgical-site infection after cardiac surgery: incidence, microbiology, and risk factors. Infect Control Hosp Epidemiol 2005 May 21;26(5):466-472. [doi: 10.1086/502569] [Medline: 15954485]
7. Lemaignen A, Birgand G, Ghodhbane W, Alkhoder S, Lolom I, Belorgey S, et al. Sternal wound infection after cardiac surgery: incidence and risk factors according to clinical presentation. Clin Microbiol Infect 2015 Jul;21(7):674.e11-674.e18 [FREE Full text] [doi: 10.1016/j.cmi.2015.03.025] [Medline: 25882356]
8. Marggraf G, Splittgerber FH, Knox M, Reidemeister JC. Mediastinitis after cardiac surgery--epidemiology and current treatment. Eur J Surg Suppl 1999 Nov 20;165(584):12-16. [doi: 10.1080/11024159950188484] [Medline: 10890226]
9. Mekontso Dessap A, Vivier E, Girou E, Brun-Buisson C, Kirsch M. Effect of time to onset on clinical features and prognosis of post-sternotomy mediastinitis. Clin Microbiol Infect 2011 Feb;17(2):292-299 [FREE Full text] [doi: 10.1111/j.1469-0691.2010.03197.x] [Medline: 20167008]
10. Lip G. Stroke and bleeding risk assessment in atrial fibrillation: when, how, and why? Eur Heart J 2013 Apr;34(14):1041-1049. [doi: 10.1093/eurheartj/ehs435] [Medline: 23257951]
11. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991 Aug;22(8):983-988. [doi: 10.1161/01.str.22.8.983] [Medline: 1866765]
12. McElroy I, Sareh S, Zhu A, Miranda G, Wu H, Nguyen M, et al. Use of digital health kits to reduce readmission after cardiac surgery. J Surg Res 2016 Jul;204(1):1-7. [doi: 10.1016/j.jss.2016.04.028] [Medline: 27451860]
13. Mistiaen P, Poot E. Telephone follow-up, initiated by a hospital-based health professional, for postdischarge problems in patients discharged from hospital to home. Cochrane Database Syst Rev 2006 Oct 18(4):CD004510 [FREE Full text] [doi: 10.1002/14651858.CD004510.pub3] [Medline: 17054207]
14. Lushaj EB, Nelson K, Amond K, Kenny E, Badami A, Anagnostopoulos PV. Timely Post-discharge Telephone Follow-Up is a Useful Tool in Identifying Post-discharge Complications Patients After Congenital Heart Surgery. Pediatr Cardiol 2016 Aug;37(6):1106-1110. [doi: 10.1007/s00246-016-1398-3] [Medline: 27064092]
15. Frias J, Virdi N, Raja P, Kim Y, Savage G, Osterberg L. Effectiveness of Digital Medicines to Improve Clinical Outcomes in Patients with Uncontrolled Hypertension and Type 2 Diabetes: Prospective, Open-Label, Cluster-Randomized Pilot Clinical Trial. J Med Internet Res 2017 Jul 11;19(7):e246 [FREE Full text] [doi: 10.2196/jmir.7833] [Medline: 28698169]
16. AliveCor. URL: https://www.alivecor.com [accessed 2020-03-17]
17. Personal MedSystems GmbH. CardioSecur. CardioSecur Research URL: https://mobile-ecg.com/research/ [accessed 2020-03-17]
18. Masimo. MightySat™ Rx Fingertip Pulse Oximeter URL: https://www.masimo.com/mightysatrx/ [accessed 2020-03-17]
19. Withings. URL: https://www.withings.com/ [accessed 2020-03-17]

20. Carroll M, Cullen T, Ferguson S, Hogge N, Horton M, Kokesh J. Innovation in Indian healthcare: using health information technology to achieve health equity for American Indian and Alaska Native populations. Perspect Health Inf Manag 2011 Jan 01;8:1d. [Medline: 21307987]
21. Zhai Y, Zhu W, Cai Y, Sun D, Zhao J. Clinical- and Cost-effectiveness of Telemedicine in Type 2 Diabetes Mellitus. Medicine 2014;93(28):e312. [doi: 10.1097/md.0000000000000312]
22. Treskes RW, van Winden LA, van Keulen N, Atsma DE, van der Velde ET, van den Akker-van Marle E, et al. Using Smart Technology to Improve Outcomes in Myocardial Infarction Patients: Rationale and Design of a Protocol for a Randomized Controlled Trial, The Box. JMIR Res Protoc 2017 Sep 22;6(9):e186 [FREE Full text] [doi: 10.2196/resprot.8038] [Medline: 28939546]
23. Treskes RW, Wolterbeek R, van der Velde ET, Eindhoven DC, Schalij MJ. Comparison of the diagnostic accuracy of four smartphone-compatible blood pressure monitors in post-myocardial infarction patients. J Telemed Telecare 2017 Apr 29;24(6):404-409. [doi: 10.1177/1357633x17704092]
24. Treskes RW, Gielen W, Wermer MJ, Grauss RW, van Alem AP, Dehnavi RA, et al. Mobile phones in cryptogenic strOke patients Bringing sIngle Lead ECGs for Atrial Fibrillation detection (MOBILE-AF): study protocol for a randomised controlled trial. Trials 2017 Aug 29;18(1):402 [FREE Full text] [doi: 10.1186/s13063-017-2131-0] [Medline: 28851409]
25. Lau JK, Lowres N, Neubeck L, Brieger DB, Sy RW, Galloway CD, et al. iPhone ECG application for community screening to detect silent atrial fibrillation: a novel technology to prevent stroke. Int J Cardiol 2013 Apr 30;165(1):193-194. [doi: 10.1016/j.ijcard.2013.01.220] [Medline: 23465249]
26. Chan P, Wong C, Poh YC, Pun L, Leung WW, Wong Y, et al. Diagnostic Performance of a Smartphone ‐ Based Photoplethysmographic Application for Atrial Fibrillation Screening in a Primary Care Setting. JAHA 2016 Jul 06;5(7). [doi: 10.1161/jaha.116.003428]
27. Haberman ZC, Jahn RT, Bose R, Tun H, Shinbane JS, Doshi RN, et al. Wireless Smartphone ECG Enables Large-Scale Screening in Diverse Populations. J Cardiovasc Electrophysiol 2015 May;26(5):520-526. [doi: 10.1111/jce.12634] [Medline: 25651872]
28. Dower GE, Yakush A, Nazzal SB, Jutzy RV, Ruiz CE. Deriving the 12-lead electrocardiogram from four (EASI) electrodes. Journal of Electrocardiology 1988 Jan;21:S182-S187. [doi: 10.1016/0022-0736(88)90090-8]
29. Sanna T, Diener H, Passman RS, Di Lazzaro V, Bernstein RA, Morillo CA, et al. Cryptogenic Stroke and Underlying Atrial Fibrillation. N Engl J Med 2014 Jun 26;370(26):2478-2486. [doi: 10.1056/nejmoa1313600]
30. Prinable JB, Foster JM, McEwan AL, Young PM, Tovey E, Thamrin C. Motivations and Key Features for a Wearable Device for Continuous Monitoring of Breathing: A Web-Based Survey. JMIR Biomed Eng 2017 Jun 26;2(1):e1. [doi: 10.2196/biomedeng.7143]
31. Adams KF, Zannad F. Clinical definition and epidemiology of advanced heart failure. American Heart Journal 1998 Jun;135(6):S204-S215. [doi: 10.1016/s0002-8703(98)70251-0]
32. Easton JD, Saver JL, Albers GW, Alberts MJ, Chaturvedi S, Feldmann E, et al. Definition and Evaluation of Transient Ischemic Attack. Stroke 2009 Jun;40(6):2276-2293. [doi: 10.1161/strokeaha.108.192218]
33. SPRINT Research Group, Wright JT, Williamson JD, Whelton PK, Snyder JK, Sink KM, et al. A Randomized Trial of Intensive versus Standard Blood-Pressure Control. N Engl J Med 2015 Nov 26;373(22):2103-2116 [FREE Full text] [doi: 10.1056/NEJMoa1511939] [Medline: 26551272]
34. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res 2013 Sep;22(7):1717-1727 [FREE Full text] [doi: 10.1007/s11136-012-0322-4] [Medline: 23184421]
35. Lowres N, Mulcahy G, Gallagher R, Ben Freedman S, Marshman D, Kirkness A, et al. Self-monitoring for atrial fibrillation recurrence in the discharge period post-cardiac surgery using an iPhone electrocardiogram. Eur J Cardiothorac Surg 2016 Jul;50(1):44-51. [doi: 10.1093/ejcts/ezv486] [Medline: 26850266]
36. R Core Team. R Project. Vienna, Austria: R Foundation for Statistical Computing The R Project for Statistical Computing URL: http://www.r-project.org [accessed 2020-03-17]
37. Stevenson LW, Pagani FD, Young JB, Jessup M, Miller L, Kormos RL, et al. INTERMACS profiles of advanced heart failure: the current picture. J Heart Lung Transplant 2009 Jun;28(6):535-541. [doi: 10.1016/j.healun.2009.02.015] [Medline: 19481012]
38. Kirklin JK, Naftel DC, Stevenson LW, Kormos RL, Pagani FD, Miller MA, et al. INTERMACS database for durable devices for circulatory support: first annual report. J Heart Lung Transplant 2008 Oct;27(10):1065-1072. [doi: 10.1016/j.healun.2008.07.021] [Medline: 18926395]
39. Pagani FD, Stevenson LW, Kirklin JK, Naftel DC, Kormos RL, Young JB. Abstract 2095: Influence of Patient Profiles on Outcomes With Implantable Mechanical Circulatory Support (MCS): Results from the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS). Circulation 2007;116:454-454. [doi: 10.1016/b978-1-4160-6001-7.00024-5]
40. Patrat-Delon S, Rouxel A, Gacouin A, Revest M, Flécher E, Fouquet O, et al. EuroSCORE II underestimates mortality after cardiac surgery for infective endocarditis. Eur J Cardiothorac Surg 2016 Mar;49(3):944-951. [doi: 10.1093/ejcts/ezv223] [Medline: 26116921]

41. Wang T, Wang M, Pemberton J. Risk Scores Predicting Mortality in Surgery of Infective Endocarditis: A Meta-Analysis. Heart, Lung and Circulation 2016 Aug;25:S296-S297. [doi: 10.1016/j.hlc.2016.06.698]
42. Huang S, Huang S, Wang C, Wu I, Chi N, Yu H, et al. Risk factors and outcome analysis after surgical management of ventricular septal rupture complicating acute myocardial infarction: a retrospective analysis. J Cardiothorac Surg 2015 May 04;10:66 [FREE Full text] [doi: 10.1186/s13019-015-0265-2] [Medline: 25935413]
43. Takahashi H, Arif R, Almashhoor A, Ruhparwar A, Karck M, Kallenbach K. Long-term results after surgical treatment of postinfarction ventricular septal rupture. European Journal of Cardio-Thoracic Surgery 2014 Jul 03;47(4):720-724. [doi: 10.1093/ejcts/ezu248]
44. Wallinder A, Pellegrino V, Fraser JF, McGiffin DC. ECMO as a bridge to non-transplant cardiac surgery. J Card Surg 2017 Aug;32(8):514-521. [doi: 10.1111/jocs.13172] [Medline: 28672423]
45. Centraal Bureau voor de Statistiek (CBS). Nederland koploper in Europa met internettoegang URL: https://www.cbs.nl/nl-nl/nieuws/2018/05/nederland-koploper-in-europa-met-internettoegang [accessed 2020-03-17]
46. Centraal Bureau voor de Statistiek (CBS). 2019 Oct 08. Internet; toegang, gebruik en faciliteiten URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED/table?dl=538C [accessed 2020-03-17]
47. Glotzer TV, Hellkamp AS, Zimmerman J, Sweeney MO, Yee R, Marinchak R, et al. Atrial High Rate Episodes Detected by Pacemaker Diagnostics Predict Death and Stroke. Circulation 2003 Apr;107(12):1614-1619. [doi: 10.1161/01.cir.0000057981.70380.45]
48. Healey JS, Connolly SJ, Gold MR, Israel CW, Van Gelder IC, Capucci A, ASSERT Investigators. Subclinical atrial fibrillation and the risk of stroke. N Engl J Med 2012 Jan 12;366(2):120-129. [doi: 10.1056/NEJMoa1105575] [Medline: 22236222]
49. Swiryn S, Orlov MV, Benditt DG, DiMarco JP, Lloyd-Jones DM, Karst E, et al. Clinical Implications of Brief Device-Detected Atrial Tachyarrhythmias in a Cardiac Rhythm Management Device Population. Circulation 2016 Oct 18;134(16):1130-1140. [doi: 10.1161/circulationaha.115.020252]
50. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur J Cardiothorac Surg 2016 Nov;50(5):e1-e88. [doi: 10.1093/ejcts/ezw313] [Medline: 27663299]

Abbreviations

AF: atrial fibrillation
BP: blood pressure
ECG: electrocardiogram
ECMO: extracorporeal membrane oxygenation
EQ-5D-5L: five-level EuroQol five-dimension
IABP: intra-aortic balloon pump
INTERMACS: Interagency Registry for Mechanically Assisted Circulatory Support
LUMC: Leiden University Medical Center
MACE: major adverse cardiac event
mHealth: mobile health
NP: nurse practitioner
POAF: postoperative atrial fibrillation

Edited by G Eysenbach; submitted 19.09.19; peer-reviewed by H Mufti, H Ross; comments to author 12.11.19; revised version received 19.12.19; accepted 26.02.20; published 21.04.20.

Please cite as:
Biersteker TE, Boogers MJ, de Lind van Wijngaarden RAF, Groenwold RHH, Trines SA, van Alem AP, Kirchhof CJHJ, van Hof N, Klautz RJM, Schalij MJ, Treskes RW
Use of Smart Technology for the Early Diagnosis of Complications After Cardiac Surgery: The Box 2.0 Study Protocol
JMIR Res Protoc 2020;9(4):e16326
URL: http://www.researchprotocols.org/2020/4/e16326/
doi:10.2196/16326
PMID:32314974

Protocol

Geriatric Trauma – A Rising Tide. Assessing Patient Safety Challenges in a Vulnerable Population Using Norwegian Trauma Registry Data and Focus Group Interviews: Protocol for a Mixed Methods Study

Mathias Cuevas-Østrem[1,2,3], MD; Olav Røise[2,3,4], MD, PhD; Torben Wisborg[5,6], MD, PhD; Elisabeth Jeppesen[1,2,3], PhD

[1]Department of Research, Norwegian Air Ambulance Foundation, Oslo, Norway

[2]Norwegian Trauma Registry, Division of Orthopaedic Surgery, Oslo University Hospital, Oslo, Norway

[3]Faculty of Health Sciences, University of Stavanger, Stavanger, Norway

[4]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[5]Norwegian National Advisory Unit on Trauma, Division of Emergencies and Critical Care, Oslo University Hospital, Oslo, Norway

[6]Anaesthesia and Critical Care Research Group, Faculty of Health Sciences, University of Tromsø - The Arctic University of Norway, Tromsø, Norway

Corresponding Author:
Mathias Cuevas-Østrem, MD
Department of Research
Norwegian Air Ambulance Foundation
Postboks 414 Sentrum
Oslo, 0103
Norway
Email: mathias.cuevas-ostrem@norskluftambulanse.no

Abstract

Background: Elderly trauma patients constitute a vulnerable group, with a substantial risk of morbidity and mortality even after low-energy falls. As the world's elderly population continues to increase, the number of elderly trauma patients is expected to increase. Limited data are available about the possible patient safety challenges that elderly trauma patients face. The outcomes and characteristics of the Norwegian geriatric trauma population are not described on a national level.

Objective: The aim of this project is to investigate whether patient safety challenges exist for geriatric trauma patients in Norway. An important objective of the study is to identify risk areas that will facilitate further work to safeguard and promote quality and safety in the Norwegian trauma system.

Methods: This is a population-based mixed methods project divided into 4 parts: 3 quantitative retrospective cohort studies and 1 qualitative interview study. The quantitative studies will compare adult (aged 16-64 years) and elderly (aged ≥65 years) trauma patients captured in the Norwegian Trauma Registry (NTR) with a date of injury from January 1, 2015, to December 31, 2018. Descriptive statistics and relevant statistical methods to compare groups will be applied. The qualitative study will comprise focus group interviews with doctors responsible for trauma care, and data will be analyzed using a thematic analysis to identify important themes.

Results: The project received funding in January 2019 and was approved by the Oslo University Hospital data protection officer (No. 19/16593). Registry data have been extracted for 33,344 patients, and the analysis of these data has begun. Focus group interviews will be conducted from spring 2020. Results from this project are expected to be ready for publication from fall 2020.

Conclusions: By combining data from the NTR with interviews with doctors responsible for treatment and transfer of elderly trauma patients, we will provide increased knowledge about trauma in Norwegian geriatric patients on a national level that will form the basis for further research aiming at developing interventions that hopefully will make the trauma system better equipped to manage the rising tide of geriatric trauma.

International Registered Report Identifier (IRRID): PRR1-10.2196/15722

(JMIR Res Protoc 2020;9(4):e15722) doi:10.2196/15722

KEYWORDS

major trauma; multiple trauma; aging; older adults; elderly; brain injuries, traumatic; geriatric; epidemiology; trauma registries; quality of health care; injury severity score

Introduction

Background

Many high- and middle-income countries around the world face the same demographic changes: people are living longer, birth rates are decreasing, and, consequently, elderly people constitute a rapidly growing proportion of the population [1,2]. The elderly often live independent and active lives despite chronic diseases and frailty and can sustain severe injury from even minor trauma [3-5]. Statistics Norway estimates that within 15 years, more people living in Norway will be aged above 65 years than below 20 years, for the first time [6]. The same report projects that by 2060, the number of Norwegians aged above 70 years will be more than double compared with the number in 2018 [6]. Consequently, there is an increase in the number of geriatric trauma patients, and the geriatric trauma population has been described as a rising tide [7].

Trauma is one of the leading causes of mortality and morbidity worldwide and in all age groups [8,9]. In Norway in 2016, the most common injuries across all ages occurred in the extremities (38.3%), head (35.4%), chest (29.5%), and spine (24.3%) [10]. Geriatric trauma patients have higher mortality rates than younger patients, adjusted for the same severity of trauma, and head injury is the leading cause of death [11-13]. Risk factors associated with a poor outcome for this group include age, pre-existing medical conditions, anticoagulant use, frailty, and altered physiological response to trauma [14-20]. Hence, geriatric trauma patients are a vulnerable group.

There is an evident shift in the epidemiology of major trauma: what used to be the disease of young men in high-energy accidents is now becoming the disease of elderly patients, where the primary mechanism of injury (MOI) is falling from less than 2 meters [21,22]. Major trauma is usually defined using the Injury Severity Score (ISS) or New Injury Severity Score (NISS), and the most common threshold is ISS >15 [23]. It has been questioned if this is too high for geriatric trauma patients, as the frail elderly might have significant morbidity and mortality even at low thresholds [24]. The age of 65 years is widely used as a cutoff for defining geriatric trauma [16,22,25-27].

Characteristics of the Geriatric Trauma Population

A 2017 report from the UK Trauma Audit and Research Network gives new and thorough insight into the characteristics of geriatric major trauma patients [22]. Some of the central findings were that over 60% of trauma patients aged 70 years and above are injured indoors, the head was the most commonly injured body region, older people admitted to hospitals had a low trauma team activation rate, and the grade of the most senior clinician treating the patients on arrival decreased with increasing patient age [22]. Low-energy trauma attracts little attention.

A geriatric trauma patient is not simply an injured old adult. Pharmacological and age-related physiological alterations in different organ systems affect the way the geriatric patient responds to both disease and injury [28]. Among the changes relevant for trauma care is that geriatric patients are often frail, meaning they have low physiological reserves [14]; they present with a higher Glasgow Coma Scale (GCS) score compared with younger patients with the same injury severity [29]; the threshold for hypotension is suggested to be 110 mm Hg, not 90 mm Hg [30,31]; and with increased age, the use of physiology-altering medications such as beta antagonists or anticoagulants increases. This might mask the severity of injury as the vital signs resemble what is considered to be within normal range values for adults. As a consequence, an injured elderly patient might seem less injured when standard triage tools are used, and this is reflected in the high rate of undertriage for geriatric major trauma patients [3,32,33]. Undertriage increases the risk of not being treated at the right level of care at the right time and can, subsequently, increase the risk of mortality [32].

Major trauma is a time-critical event; hence, disposing the right resources at the right time without unjustifiable delay is crucial. Paradoxically, it is the elderly patients—the ones with the least physiological reserves—who get delayed treatment [22,34,35]. Both Advanced Trauma Life Support (ATLS) and the Eastern Association for the Surgery of Trauma geriatric trauma guidelines advocate for an aggressive treatment approach until otherwise decided [27,28]. Early and aggressive treatment is shown to increase survival rates in older trauma patients [36].

Traumatic Brain Injury

Traumatic brain injury (TBI) is one of the leading causes of trauma-related deaths [37]. Antiplatelet and anticoagulant drugs are frequently used in the geriatric trauma population, a risk factor for acute intracranial bleeding following head injury. A computed tomography (CT) head scan is needed to detect bleedings, and this can be done in all acute care trauma hospitals in Norway. In cases of moderate-to-severe TBI, the acute care trauma hospital can contact the neurosurgical department in the regional trauma center for clinical guidance and assessment of patient transfer. Experience from clinical practice nationally and internationally shows that the transfer of elderly trauma patients with head injury to a neurosurgical facility from an acute care trauma hospital is a challenge 35. We believe that there are more factors than just injury, severity, and national transfer criteria that determine whether patients are transferred from an acute care trauma hospital to a trauma center with neurosurgical facilities. We believe that possible factors are age, comorbidities, activities of daily life functions, prognosis, limitations in ward capacity, limitations in what the neurosurgical intervention can offer to improve prognosis, and limited time before it is too late to intervene, along with culture and an expectation of a negative outcome. This will be explored in this project.

Norwegian Trauma System

The 2016 National Trauma Plan for Norway provides requirements for all services in the national trauma system—from prehospital care to rehabilitation [38]. Norway has 2 hospital levels treating trauma patients; 34 acute care trauma hospitals and 4 trauma centers. Acute care trauma hospitals are spread out around the country, and trauma centers are regional university hospitals. All acute care trauma hospitals offer general surgical and orthopedic services and are capable of stabilizing severely injured patients before transferring them to trauma centers, if necessary, but do not offer neurosurgery, intervention radiology (except for a few), and other specialized services. The trauma centers offer all medical specialties, including neurosurgery, and are capable of managing all types of injuries [39]. The annual number of patients meeting the inclusion criteria of the Norwegian Trauma Registry (NTR; see Methods) is approximately 8000 [40].

NTR is a national medical quality registry that has been operational from January 2015. The objective of the registry is to monitor trauma treatment in Norway and to contribute to increased treatment quality. All acute care trauma hospitals and trauma centers in Norway report to the registry. These hospitals have certified registrars who register data from injury to rehabilitation after the Utstein template and classify all injuries according to the Abbreviated Injury Scale (AIS) and calculate ISS and NISS [41,42]. All patients receive written information about the registry, including the opportunity to access the data recorded and to deny registration.

Aims and Objectives

The aims of this project are to investigate whether patient safety challenges exist for elderly trauma patients in Norway and to identify risk areas that will facilitate further work to safeguard and promote quality and safety in the Norwegian trauma system. A total of 3 retrospective cohort studies and 1 qualitative interview study will be conducted. The results of each study will be published in peer-reviewed medical journals.

The specific objectives of the project are as follows:

1. To assess whether injured elderly Norwegian patients (65 years) are given different emergency trauma care compared with younger patients.
2. To explore explanations for potential differences in the quality of trauma care between age groups in the emergency part of the trauma chain.

The quantitative studies aim to achieve the following:

- Determine the characteristics of geriatric trauma patients in Norway and compare this group with the Norwegian adult population and results from comparable international publications.
- Describe differences between the adult and elderly general and TBI populations in Norway regarding injury severity, MOI, 30-day mortality, hospital level of care, transport methods, emergency interventions, radiological examinations, and physiological variables.

The qualitative interview study aims to achieve the following:

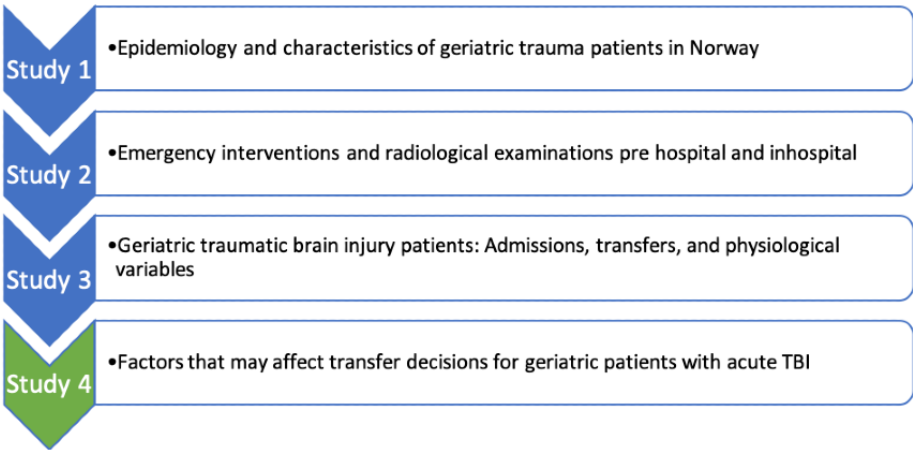
- Explore factors that may affect transfer decisions for geriatric patients with TBI.

Methods

Study Design

The *Geriatric Trauma—Assessing Patient Safety* project applies a mixed methods design, and it consists of 3 quantitative retrospective cohort studies using data from the NTR and 1 qualitative interview study focusing on the management of patients with acute TBI. The 4 studies included in the project are presented in Figure 1. The use of both qualitative and quantitative methods provides a deeper understanding of the processes involved in the care of elderly TBI patients and can increase the understanding of causative factors regarding the management of this group. The qualitative study provides an extra layer of information that will help interpret the quantitative data on TBI so that it can be better used in improving the trauma system.

Figure 1. An overview of the four studies included in the project. TBI: traumatic brain injury.



Study Setting

The main study setting is the prehospital and emergency department part of the Norwegian trauma system. Data from the NTR are collected from all acute care trauma hospitals and trauma centers in Norway, which has a population of about 5.3 million inhabitants spread out over vast distances with a mix of urban and rural areas. For the qualitative study, we seek to include a sample with representatives from at least two Norwegian health regions.

Study Participants: Data From the Norwegian Trauma Registry

The NTR was searched to identify all trauma patients included in the registry from January 1, 2015, to December 31, 2018. A total of 33, 344 patients were included in the registry, of which 22,603 were aged between 16 and 64 years and 6334 were 65 years or older. A total of 3735 elderly patients had NISS above or equal to 9. The eligibility criteria for the registry are presented in Textboxes 1 and 2.

Textbox 1. Inclusion criteria for the Norwegian Trauma Registry.

- All patients admitted with trauma team activation (TTA) on arrival to the emergency department in all acute care trauma hospitals and trauma centers in Norway, irrespective of Injury Severity Score and New Injury Severity Score (NISS)
- All patients treated at an acute care trauma hospital or trauma center in Norway, without TTA, with one or more of the following injuries:
  - Penetrating injury to the head, neck, torso, or extremities proximal to the elbow or knee
  - Head injury with abbreviated injury score (AIS) ≥3
  - NISS >12
- All patients with trauma-related deaths at the site of trauma or during transportation to the hospital, who are not referred to the hospital, but where prehospital management or treatment is initiated

Textbox 2. Exclusion criteria for the Norwegian Trauma Registry.

- Patients with chronic subdural hematoma, without other trauma-related injuries
- Patients with injuries from drowning, inhalation, hypothermia, and asphyxia without concomitant trauma
- Patients who die on scene without the activation of prehospital resources

Quantitative Registry-Based Retrospective Cohort Studies

The 3 quantitative studies are all registry-based retrospective cohort studies focusing on (1) epidemiology and characteristics, (2) emergency interventions and radiological examinations, and (3) TBI in the Norwegian geriatric trauma cohort. The specific outcome measures for each study are presented in Table 1, and a full overview of the variables extracted from the registry is presented in Multimedia Appendix 1.

Study number three focuses on elderly TBI patients, a particularly vulnerable patient group with high morbidity and mortality [37]. The severity of TBI can be defined using different measures. AIS is an international classification system defining all injury types according to severity where 1 is minor and 6 is maximal and currently untreatable [42]. AIS ≥ 3 is recognized as moderate-to-severe head injury.. GCS at presentation is one of the major factors directing neurosurgical decision making, traditionally classifying TBI into mild (GCS 13-15), moderate (GCS 9-12), and severe (GCS <8). Recent evidence suggests that GCS is not as sensitive for detecting TBI in the elderly, so we will do analyses for both parameters [29]. In addition, GCS is the only measure of the two with prehospital value. Patients admitted with a low GCS not caused by head trauma will be excluded from these analyses.

Table 1. Overview of the quantitative studies.

Characteristics	Study 1	Study 2	Study 3
Aims	• Describe the Norwegian geriatric trauma population and assess differences in demographic and epidemiological characteristics between age groups • Assess 30-day mortality • Identify injury mechanism differences between age groups • Assess differences in the prehospital and inhospital levels of care between age groups	• Assess differences in the proportion of emergency interventions (prehospital and inhospital airway management and pneumothorax decompression) and radiological examinations (inhospital) performed on elderly and younger patients • Assess differences in emergency interventions and radiological examinations performed on elderly and younger patients on the basis of clinical findings	• Assess differences in admission rates and transfer rates to trauma centers with neurosurgical services for patients in different age groups with moderate-to-severe traumatic brain injury • Assess differences between age groups in transport method (car or air ambulance) for patients with same degree of injury severity • Assess differences in physiological variables between age groups, both prehospital and at admission (systolic blood pressure, respiratory rate, Glasgow Coma Scale score, and body temperature) for patients with the same degree of injury severity
Hypothesis	• Younger patients suffer primarily from injury due to high-energy trauma, and elderly patients suffer primarily from injury due to low-energy trauma • Younger patients have higher admission rates to trauma centers than the elderly for similar injury severity • Younger patients have higher transfer rates to trauma centers than the elderly for similar injury severity	• Prehospital personnel use the same algorithm in decision making in both elderly and younger patients, that is, there is no discrimination in how elderly and younger patients with the same vital signs are treated • The elderly population is expected to have same frequencies of examinations and interventions as the younger patients for the same severity of injuries, both prehospital and in the emergency room	• Younger patients have higher admission rates to trauma centers than the elderly • Younger patients have higher transfer rates to trauma centers than the elderly • Younger patients are more often transported by air ambulance than the elderly
Outcome measures	Primary: • 30-day mortality Secondary: • Age, gender, mechanism of injury, blunt or penetrating trauma, Abbreviated Injury Scale, Injury Severity Scale, New Injury Severity Scale • Location of injury • Time from injury to admission • Transport method • Level of prehospital and inhospital care • Interventions given prehospitally and in the emergency department • Trauma team activation • Level of care at admission and discharge • Length of stay	Primary: • Number and type of radiological examinations and emergency interventions (frequencies) Secondary: • Time to examination (x-ray; thorax, pelvis, and computed tomography) • Physiological variables	Primary: • 30-day mortality Secondary: • Admissions to acute care trauma hospitals and trauma centers • Transfers to higher level of care • Transport methods • Physiological variables • Interventions given prehospitally and in the emergency department

Data Analysis

All injured adult patients admitted to a Norwegian hospital and registered in the NTR in the period January 1, 2015, to December 31, 2018, will be included in the analysis. Trauma registry data will be analyzed using descriptive statistical methods and other relevant statistical methods to compare adult (aged 16-64 years) with elderly (aged 65 years) trauma patients, as described below. Data might also be analyzed to compare subgroups, for example, 10-year age intervals, if the data allow it. Data will be reported following the Strengthening the Reporting of Observational Studies in Epidemiology statement

checklist. Categorical variables will be analyzed using a Pearson chi-square test, continuous variables will be analyzed with normal score distribution using *t* tests, and skewed distributions will be analyzed using the Mann-Whitney *U* test. We will consider using the Fisher exact test for smaller subgroups. We will also consider doing a logistic regression analysis. The strength of association will be reported as an odds ratio with 95% CI. Low statistical power because of small groups and few events could result in some significant differences with broad 95% CIs. The correlation between the age groups is planned to be tested with a Spearman rank correlation test. We consider our study as explorative, and the significance level will therefore be kept at $P<.05$. The analyses would be performed by using SPSS version 25 or higher (IBM SPSS Statistics for Mac, IBM Corporation).

All data will be handled and saved in a secured data server administered by the Norwegian Air Ambulance Foundation. All data will be unidentifiable when sharing between the authors and in the analysis and presentations. Data will be stored for 5 years after the project is finished.

Study 4: Qualitative Interview Study Addressing Factors That May Affect and Explain Transfer Decisions for Geriatric Patients With Acute Traumatic Brain Injury

Participants

A sample of participants for the focus group interviews will be recruited from doctors responsible for the treatment and transfer decisions for head trauma patients. We seek to include participants with the following characteristics:

- Acute care trauma hospital team leaders: Responsible for initial evaluation, transfer evaluation, and continued care in case the patient is not transferred. We seek registrars or consultants with more than 1 year of experience as a trauma team leader and trained in the ATLS principles according to the requirements in the national trauma plan. The subjects should preferably have experienced at least one case of a geriatric trauma patient where head trauma was the main reason for discussing transfer.
- Neurosurgeons in trauma centers: Taking part in decision making on accepting the patient for transfer or not, being responsible for all neurosurgical interventions, monitoring, and care in a neurosurgical ward. We seek registrars or consultants with more than 1 year of experience in on-call decision making, assessing patients for transfer to their respective hospitals.

A priori, it is estimated that 4 focus group interviews will be sufficient, but data acquisition will continue until saturation is reached. The interviews will be conducted separately (mono-professional) to reveal possible professional differences. The groups will be recruited using a combination of the snowball sampling method to reach out to a wide network, purposive sampling to include doctors with first-hand experience with geriatric head trauma and working locations in different health regions, and convenience sampling to conduct interviews in regional or national forums. All participants will receive written and oral information about the purpose of the study. We will also obtain informed consent. Before starting the interview, they will be informed that they are discussing factors affecting management and transfer decisions in patients with TBI. The interviewer will use an interview guide with open-ended questions to ensure that the relevant subjects are covered. This will cover themes such as priorities and ethical considerations, patient-related factors emphasized in the decision-making process, guidelines, attitudes, culture, and interventions.

Analysis

The interviews will be audio-recorded and transcribed verbatim. The data found in the interviews will be categorized and analyzed using thematic analysis as described by Braun and Clark [43]. Each interview will be coded by at least two analysts who will read the transcripts and, if appropriate, listen to the audio recordings to ensure the proper meaning is captured. The analysts will generate codes and sort these into themes. Coding disparities or uncertainties will be discussed with additional researchers in the group.

Ethical Considerations

Research will be conducted according to the ethical guidelines of the Helsinki declaration. The study protocol is approved by the Oslo University Hospital data protection officer, which is responsible for the Norwegian Trauma Registry (No. 19/16593).

Results

Registry data have been extracted on 33,344 patients, and the analysis of data for study number one is ready to be performed. Anticipated findings are that the Norwegian geriatric trauma population shows a number of similar characteristics as described in papers from comparable Western populations (e.g. the Netherlands, United Kingdom, and Australia) but that the proportion of geriatric trauma patients is smaller than that in countries with a larger elderly population, for example, Japan [22,44-46]. The next steps will be to work parallelly on the manuscript on study 1 and on conducting interviews for study 4. Studies 2 and 3 will be conducted subsequently. The project plan has been presented in relevant forums in Norway and Europe. Results from each study will be published in peer-reviewed medical journals from 2020.

Discussion

Principal Considerations

The vulnerable population of geriatric trauma patients is increasing in number. It is a group with clinically challenging characteristics, such as comorbidity, polypharmacy and frailty, and a high risk of undertriage. As major trauma shifts from being a disease of the young to a disease of the elderly injured in low-energy accidents, substantial patient safety risks may exist, for example, differences in the level of care between adults and elderly patients. To our knowledge, no study has been conducted in Norway by using national data assessing such differences in trauma care.

Kirkman et al [35] published a paper that raises the following central question: “Do elderly head injuries do worse because of a self-fulfilling prophecy of poorer management?” They

found that the time from admission to CT head imaging and the likelihood of not being transferred to a center with acute neurosurgical care facilities increased with age. Another study from Utter et al [34] found that geriatric trauma patients have delayed transfer to a neurocenter in a level I trauma center [34]. Little is known about which factors affect the decisions that lead to this. Negative attitudes toward elderly patients and an expectation of a poor outcome might lead to a passive, observing role and low treatment ambitions, and this will be addressed in this project.

In 2019, a paper about geriatric trauma patients from the largest trauma center in Norway was published, showing that mortality increased with age and was inversely related to the probability of trauma team activation on arrival [13]. Moreover, Australian, Dutch, British, and Japanese papers published in the recent years give a thorough overview of the characteristics of geriatric trauma populations in comparable countries [22,44-46]. Whether the total Norwegian trauma population shares some of these characteristics is not known.

Strengths and Limitations

The project employs a mixed methods design, where possible patient safety challenges of the geriatric trauma population will be assessed through 4 studies. The mixed methods design is one of the project's strengths, as the qualitative methodology brings forward information that the registry data cannot provide and makes the interpretation of the retrospective data more reliable when it should be translated into clinical practice. Another strength with this project is how it focuses particularly on the potential patient safety challenges of elderly trauma patients. As far as we know, it is the first project on geriatric trauma patients where patient safety is the overarching theme. A high generalizability to other trauma systems is expected, given the similarities between demographical changes and trauma systems in many high-income countries. Limitations are inherent to the retrospective design of the quantitative studies, with risk of bias and the fact that causal factors cannot be explored.

Conclusions

With the rising tide of geriatric trauma as a background, this research will have a societal impact. If there are differences between adult and elderly trauma patients, it is important to know to make sound decisions in the future. For example, if geriatric trauma patients are found to be systematically treated at a lower level of trauma care, it will be important to document this, and the next step will be to examine why these differences exist. Findings regarding characteristics and physiological responses will possibly support international studies, and ours will be the first study to assess this in the Norwegian population. With the rising tide of geriatric trauma fast approaching, we want to investigate differences in trauma care between age groups in the Norwegian population and evaluate if patient safety risks exist for geriatric trauma patients.

Acknowledgments

The project is a collaboration among the Norwegian Air Ambulance Foundation, the NTR, Oslo University Hospital, Norwegian National Advisory Unit on Trauma, and the University of Stavanger. The project leader is Professor OR MD, PhD. The study will be performed by one PhD fellow, MCØ.

The project has received funding from the Norwegian Air Ambulance Foundation from January 2019 to December 2021.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of extracted variables from the Norwegian Trauma Registry.
[DOCX File , 18 KB - resprot_v9i4e15722_app1.docx]

References

1. World Health Organization. 2018 Feb 5. Ageing and Health URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2019-06-04]
2. Vincent GK, Velkoff VA. Census. Washington, DC: US Department of Commerce Economics and Statistics Administration, US Census Bureau; 2010 May. The Next Four Decades: The Older Population in the United States: 2010 to 2050. Population Estimates and Projections URL: https://www.census.gov/content/dam/Census/library/publications/2010/demo/p25-1138.pdf [accessed 2019-06-04]
3. Kodadek LM, Selvarajah S, Velopulos CG, Haut ER, Haider AH. Undertriage of older trauma patients: is this a national phenomenon? J Surg Res 2015 Nov;199(1):220-229. [doi: 10.1016/j.jss.2015.05.017] [Medline: 26070496]
4. Vaage OF. Statistisk sentralbyrå (SSB) / Statistics Norway. 2015 Mar. [Over 80 and Still Physically Active] URL: https://www.ssb.no/sosiale-forhold-og-kriminalitet/ssp/_attachment/242854?_ts=150652906d8 [accessed 2019-06-04]
5. Bergeron E, Clement J, Lavoie A, Ratte S, Bamvita J, Aumont F, et al. A simple fall in the elderly: not so simple. J Trauma 2006 Feb;60(2):268-273. [doi: 10.1097/01.ta.0000197651.00482.c5] [Medline: 16508481]
6. Syse A, Leknes S, Løkken S, Tønnesen M. Norway's 2018 Population Projections. Main Results, Methods and Assumptions. Oslo–Kongsvinger: Statistics Norway; 2018.

7. Ciesla DJ, Pracht EE, Tepas JJ, Cha JY, Langland-Orban B, Flint LM. The injured elderly: a rising tide. Surgery 2013 Aug;154(2):291-298. [doi: 10.1016/j.surg.2013.04.025] [Medline: 23889955]
8. Knudsen AK, Tollånes MC, Haaland Ø, Kinge JM, Skirbekk V, Vollset SE. Disease Burden in Norway 2015. Results from the Global Burden of Diseases, Injuries, and Risk Factors Study 2015. Bergen/Oslo: Folkehelseinstituttet; 2017.
9. World Health Organization. 2010. Injuries and Violence: The Facts URL: https://www.who.int/violence_injury_prevention/key_facts/en/ [accessed 2019-06-23]
10. Jeppesen E, Hestnes M, Ringdal K, Røise O. Nasjonalt servicemiljø for medisinske kvalitetsregistre. Oslo: Oslo University Hospital; 2017. Annual report of the Norwegian Trauma Registry URL: https://www.kvalitetsregistre.no/sites/default/files/40_arsrapport_2016_norsk_traumeregister.pdf [accessed 2019-06-19]
11. Hashmi A, Ibrahim-Zada I, Rhee P, Aziz H, Fain MJ, Friese RS, et al. Predictors of mortality in geriatric trauma patients: a systematic review and meta-analysis. J Trauma Acute Care Surg 2014 Mar;76(3):894-901. [doi: 10.1097/TA.0b013e3182ab0763] [Medline: 24553567]
12. Perdue PW, Watts DD, Kaufmann CR, Trask AL. Differences in mortality between elderly and younger adult trauma patients: geriatric status increases risk of delayed death. J Trauma 1998 Oct;45(4):805-810. [doi: 10.1097/00005373-199810000-00034] [Medline: 9783625]
13. Ringen AH, Gaski IA, Rustad H, Skaga NO, Gaarder C, Naess PA. Improvement in geriatric trauma outcomes in an evolving trauma system. Trauma Surg Acute Care Open 2019;4(1):e000282 [FREE Full text] [doi: 10.1136/tsaco-2018-000282] [Medline: 31245616]
14. Joseph B, Pandit V, Zangbar B, Kulvatunyou N, Hashmi A, Green DJ, et al. Superiority of frailty over age in predicting outcomes among geriatric trauma patients: a prospective analysis. JAMA Surg 2014 Aug;149(8):766-772. [doi: 10.1001/jamasurg.2014.296] [Medline: 24920308]
15. Wang CY, Chen YC, Chien TH, Chang HY, Chen YH, Chien CY, et al. Impact of comorbidities on the prognoses of trauma patients: Analysis of a hospital-based trauma registry database. PLoS One 2018;13(3):e0194749 [FREE Full text] [doi: 10.1371/journal.pone.0194749] [Medline: 29558508]
16. Goodmanson NW, Rosengart MR, Barnato AE, Sperry JL, Peitzman AB, Marshall GT. Defining geriatric trauma: when does age make a difference? Surgery 2012 Oct;152(4):668-74; discussion 674 [FREE Full text] [doi: 10.1016/j.surg.2012.08.017] [Medline: 23021136]
17. Heffernan DS, Thakkar RK, Monaghan SF, Ravindran R, Adams CA, Kozloff MS, et al. Normal presenting vital signs are unreliable in geriatric blunt trauma victims. J Trauma 2010 Oct;69(4):813-820. [doi: 10.1097/TA.0b013e3181f41af8] [Medline: 20938267]
18. Sammy I, Lecky F, Sutton A, Leaviss J, O'Cathain A. Factors affecting mortality in older trauma patients-A systematic review and meta-analysis. Injury 2016 Jun;47(6):1170-1183. [doi: 10.1016/j.injury.2016.02.027] [Medline: 27015751]
19. Boltz MM, Podany AB, Hollenbeak CS, Armen SB. Injuries and outcomes associated with traumatic falls in the elderly population on oral anticoagulant therapy. Injury 2015 Sep;46(9):1765-1771. [doi: 10.1016/j.injury.2015.06.013] [Medline: 26117415]
20. Peck KA, Calvo RY, Schechter MS, Sise CB, Kahl JE, Shackford MC, et al. The impact of preinjury anticoagulants and prescription antiplatelet agents on outcomes in older patients with traumatic brain injury. J Trauma Acute Care Surg 2014 Feb;76(2):431-436. [doi: 10.1097/TA.0000000000000107] [Medline: 24458049]
21. Kehoe A, Smith JE, Edwards A, Yates D, Lecky F. The changing face of major trauma in the UK. Emerg Med J 2015 Dec;32(12):911-915 [FREE Full text] [doi: 10.1136/emermed-2015-205265] [Medline: 26598629]
22. Banerjee J, Baxter M, Coats T, Edwards A, Griffiths R, Kumar DS, et al. Trauma Audit & Research Network (TARN). 2017. Major Trauma in Older People URL: https://www.tarn.ac.uk/content/downloads/3793/Major%20Trauma%20in%20Older%20People%202017.pdf [accessed 2018-07-31]
23. Palmer C. Major trauma and the injury severity score--where should we set the bar? Annu Proc Assoc Adv Automot Med 2007;51:13-29 [FREE Full text] [Medline: 18184482]
24. Meagher AD, Lin A, Mandell SP, Bulger E, Newgard C. A comparison of scoring systems for predicting short- and long-term survival after trauma in older adults. Acad Emerg Med 2019 Jun;26(6):621-630. [doi: 10.1111/acem.13727] [Medline: 30884022]
25. Grossman MD, Miller D, Scaff DW, Arcona S. When is an elder old? Effect of preexisting conditions on mortality in geriatric trauma. J Trauma 2002 Feb;52(2):242-246. [doi: 10.1097/00005373-200202000-00007] [Medline: 11834982]
26. Hildebrand F, Pape H, Horst K, Andruszkow H, Kobbe P, Simon T, et al. Impact of age on the clinical outcomes of major trauma. Eur J Trauma Emerg Surg 2016 Jun;42(3):317-332. [doi: 10.1007/s00068-015-0557-1] [Medline: 26253883]
27. Calland JF, Ingraham AM, Martin N, Marshall GT, Schulman CI, Stapleton T, Eastern Association for the Surgery of Trauma. Evaluation and management of geriatric trauma: an Eastern Association for the Surgery of Trauma practice management guideline. J Trauma Acute Care Surg 2012 Nov;73(5 Suppl 4):S345-S350. [doi: 10.1097/TA.0b013e318270191f] [Medline: 23114492]
28. American College of Surgeons. Advanced Trauma Life Support (ATLS). Ninth Edition. Chicago: American College of Surgeons; 2012.

29. Kehoe A, Smith JE, Bouamra O, Edwards A, Yates D, Lecky F. Older patients with traumatic brain injury present with a higher GCS score than younger patients for a given severity of injury. Emerg Med J 2016 Jun;33(6):381-385. [doi: 10.1136/emermed-2015-205180] [Medline: 26825613]
30. Eastridge BJ, Salinas J, McManus JG, Blackburn L, Bugler EM, Cooke WH, et al. Hypotension begins at 110 mm Hg: redefining 'hypotension' with data. J Trauma 2007 Aug;63(2):291-7; discussion 297. [doi: 10.1097/TA.0b013e31809ed924] [Medline: 17693826]
31. Oyetunji TA, Chang DC, Crompton JG, Greene WR, Efron DT, Haut ER, et al. Redefining hypotension in the elderly: normotension is not reassuring. Arch Surg 2011 Jul;146(7):865-869. [doi: 10.1001/archsurg.2011.154] [Medline: 21768435]
32. Rogers A, Rogers F, Bradburn E, Krasne M, Lee J, Wu D, et al. Old and undertriaged: a lethal combination. Am Surg 2012 Jun;78(6):711-715. [Medline: 22643270]
33. Chang DC, Bass RR, Cornwell EE, Mackenzie EJ. Undertriage of elderly trauma patients to state-designated trauma centers. Arch Surg 2008 Aug;143(8):776-81; discussion 782. [doi: 10.1001/archsurg.143.8.776] [Medline: 18711038]
34. Utter GH, Victorino GP, Wisner DH. Interhospital transfer occurs more slowly for elderly acute trauma patients. J Emerg Med 2008 Nov;35(4):415-420. [doi: 10.1016/j.jemermed.2007.04.021] [Medline: 17933480]
35. Kirkman MA, Jenks T, Bouamra O, Edwards A, Yates D, Wilson MH. Increased mortality associated with cerebral contusions following trauma in the elderly: bad patients or bad management? J Neurotrauma 2013 Aug 15;30(16):1385-1390. [doi: 10.1089/neu.2013.2881] [Medline: 23441674]
36. Demetriades D, Karaiskakis M, Velmahos G, Alo K, Newton E, Murray J, et al. Effect on outcome of early intensive management of geriatric trauma patients. Br J Surg 2002 Oct;89(10):1319-1322. [doi: 10.1046/j.1365-2168.2002.02210.x] [Medline: 12296905]
37. Stein DM, Kozar RA, Livingston DH, Luchette F, Adams SD, Agrawal V, AAST Geriatric Trauma/ACS Committee. Geriatric traumatic brain injury-What we know and what we don't. J Trauma Acute Care Surg 2018 Oct;85(4):788-798. [doi: 10.1097/TA.0000000000001910] [Medline: 30256343]
38. Norwegian National Advisory Unit on Trauma. [National trauma plan - Trauma system in Norway 2016]. 2016. URL: https://traumeplan.no [accessed 2019-07-31]
39. Dehli T, Gaarder T, Christensen BJ, Vinjevoll OP, Wisborg T. Implementation of a trauma system in Norway: a national survey. Acta Anaesthesiol Scand 2015 Mar;59(3):384-391 [FREE Full text] [doi: 10.1111/aas.12467] [Medline: 25582880]
40. Jeppesen E, Ringdal KG, Hoem P, Røise O. Nasjonalt servicemiljø for medisinske kvalitetsregistre. Oslo: Oslo University Hospital; 2019. Annual report of the Norwegian Trauma Registry 2018 URL: https://www.kvalitetsregistre.no/sites/default/files/39_arsrapport_2019_nasjonalt_traumeregister.pdf [accessed 2019-11-05]
41. Ringdal KG, Coats TJ, Lefering R, di Bartolomeo S, Steen PA, Røise O, Utstein TCD expert panel. The Utstein template for uniform reporting of data following major trauma: a joint revision by SCANTEM, TARN, DGU-TR and RITG. Scand J Trauma Resusc Emerg Med 2008 Aug 28;16:7 [FREE Full text] [doi: 10.1186/1757-7241-16-7] [Medline: 18957069]
42. Gennarelli TA, Wodzin E. AIS 2005: a contemporary injury scale. Injury 2006 Dec;37(12):1083-1091. [doi: 10.1016/j.injury.2006.07.009] [Medline: 17092503]
43. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. Los Angeles: SAGE Publications Ltd; 2013.
44. Beck B, Cameron P, Lowthian J, Fitzgerald M, Judson R, Gabbe BJ. Major trauma in older persons. BJS Open 2018 Sep;2(5):310-318 [FREE Full text] [doi: 10.1002/bjs5.80] [Medline: 30263982]
45. de Vries R, Reininga IH, Pieske O, Lefering R, El Moumni M, Wendt K. Injury mechanisms, patterns and outcomes of older polytrauma patients-An analysis of the Dutch Trauma Registry. PLoS One 2018;13(1):e0190587 [FREE Full text] [doi: 10.1371/journal.pone.0190587] [Medline: 29304054]
46. Kojima M, Endo A, Shiraishi A, Otomo Y. Age-related characteristics and outcomes for patients with severe trauma: analysis of Japan's nationwide trauma registry. Ann Emerg Med 2019 Mar;73(3):281-290. [doi: 10.1016/j.annemergmed.2018.09.034] [Medline: 30447945]

Abbreviations

AIS: Abbreviated Injury Scale
ATLS: advanced trauma life support
CT: computed tomography
GCS: Glasgow Coma Scale
ISS: Injury Severity Score
MOI: mechanism of injury
NISS: New Injury Severity Score
NTR: Norwegian Trauma Registry
TBI: traumatic brain injury

Edited by G Eysenbach; submitted 13.08.19; peer-reviewed by B Smith, M Laurent, M Perez-Zepeda, Z Ma, J Barouh; comments to author 03.10.19; revised version received 27.11.19; accepted 17.12.19; published 30.04.20.

Please cite as:
Cuevas-Østrem M, Røise O, Wisborg T, Jeppesen E
Geriatric Trauma – A Rising Tide. Assessing Patient Safety Challenges in a Vulnerable Population Using Norwegian Trauma Registry Data and Focus Group Interviews: Protocol for a Mixed Methods Study
JMIR Res Protoc 2020;9(4):e15722
URL: http://www.researchprotocols.org/2020/4/e15722/
doi:10.2196/15722
PMID:32352386

Protocol

Determining the Agreement Between an Automated Respiratory Rate Counter and a Reference Standard for Detecting Symptoms of Pneumonia in Children: Protocol for a Cross-Sectional Study in Ethiopia

Charlotte Ward[1], MA, MSc; Kevin Baker[1,2], MA, MSc, PhD; Sarah Marks[1], MSc; Dawit Getachew[3], MD; Tedila Habte[3], MSc; Cindy McWhorter[4], BSc; Paul Labarre[4], MME, MBA; Jonathan Howard-Brand[4], BSc; Nathan P Miller[5], MIA, MPH, PhD; Hayalnesh Tarekegn[5], MB, BCh, BAO; Solomie Jebessa Deribessa[6], MPH, MD; Max Petzold[7], PhD; Karin Kallander[1,2,5], MSc, PhD

[1]Malaria Consortium, London, United Kingdom

[2]Department of Public Health Sciences, Karolinska Institutet, Solnavägen, Sweden

[3]Malaria Consortium, Addis Ababa, Ethiopia

[4]United Nations Children's Fund Supply Division, Copenhagen, Denmark

[5]United Nations Children's Fund Programme Division, New York, NY, United States

[6]St Paul's Hospital Millennium Medical College, Addis Ababa, Ethiopia

[7]School of Public Health and Community Medicine, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden

Corresponding Author:
Charlotte Ward, MA, MSc
Malaria Consortium
The Green House 244-254 Cambridge Heath Rd
London
United Kingdom
Phone: 44 0 20 3559 ext 6438
Email: c.ward@malariaconsortium.org

Abstract

Background: Acute respiratory infections (ARIs), primarily pneumonia, are the leading infectious cause of under-5 mortality worldwide. Manually counting respiratory rate (RR) for 60 seconds using an ARI timer is commonly practiced by community health workers to detect fast breathing, an important sign of pneumonia. However, correctly counting breaths manually and classifying the RR is challenging, often leading to inappropriate treatment. A potential solution is to introduce RR counters, which count and classify RR automatically.

Objective: This study aims to determine how the RR count of an Automated Respiratory Infection Diagnostic Aid (ARIDA) agrees with the count of an expert panel of pediatricians counting RR by reviewing a video of the child's chest for 60 seconds (reference standard), for children aged younger than 5 years with cough and/or difficult breathing.

Methods: A cross-sectional study aiming to enroll 290 children aged 0 to 59 months presenting to pediatric in- and outpatient departments at a teaching hospital in Addis Ababa, Ethiopia, was conducted. Enrollment occurred between April and May 2017. Once enrolled, children participated in at least one of three types of RR evaluations: (1) agreement—measure the RR count of an ARIDA in comparison with the reference standard, (2) consistency—measure the agreement between two ARIDA devices strapped to one child, and (3) RR fluctuation—measure RR count variability over time after ARIDA attachment as measured by a manual count. The agreement and consistency of expert clinicians (ECs) counting RR for the same child with the Mark 2 ARI timer for 60 seconds was also measured in comparison with the reference standard.

Results: Primary outcomes were (1) mean difference between the ARIDA and reference standard RR count (agreement) and (2) mean difference between RR counts obtained by two ARIDA devices started simultaneously (consistency).

Conclusions: Study strengths included the design allowing for comparison between both ARIDA and the EC with the reference standard RR count. A limitation is that exactly the same set of breaths were not compared between ARIDA and the reference standard since ARIDA can take longer than 60 seconds to count RR. Also, manual RR counting, even when aided by a video of

the child's chest movements, is subject to human error and can result in low interrater reliability. Further work is needed to reach global consensus on the most appropriate reference standard and an acceptable level of agreement to provide ministries of health with evidence to make an informed decision on whether to scale up new automated RR counters.

Trial Registration: ClinicalTrials.gov NCT03067558; https://clinicaltrials.gov/ct2/show/NCT03067558

International Registered Report Identifier (IRRID): RR1-10.2196/16531

(JMIR Res Protoc 2020;9(4):e16531) doi:10.2196/16531

KEYWORDS

pneumonia; diagnostics; child; respiratory rate; Ethiopia

Introduction

Acute respiratory infections (ARIs), primarily pneumonia, are the leading infectious causes of death among children aged younger than 5 years globally, accounting for an estimated 0.9 million deaths in 2015 [1], with over 75% of these deaths clustering in sub-Saharan Africa and Southeast Asia. Deaths from pneumonia in children result mostly from delayed presentation to appropriate health care providers and inappropriate treatment [2].

Diagnosis of pneumonia by community health workers (CHWs) and first-level health facility workers (FLHFWs), collectively known as frontline health workers, is based on counting the number of breaths in 60 seconds in children aged younger than 5 years with cough and/or difficulty breathing to assess whether the respiratory rate (RR) is high enough for a particular age to prescribe antibiotics and treat suspected pneumonia. This is defined by the World Health Organization (WHO) Integrated Management of Childhood Illness (IMCI) guidelines [3] for FLHFWs and the Integrated Community Case Management (iCCM) guidelines for CHWs [4]. Current standard practice for frontline health workers is to count RR manually by observing chest movements for 60 seconds. In practice defining a breath and counting RR can be difficult, as children breathe irregularly and faster than adults and the child may not be calm and still for a full minute. Misclassification of the observed rate remains high [5,6], often leading to inappropriate treatment [7].

The Acute Respiratory Infection Diagnostic Aids (ARIDA) project [8] was initiated as a response to the call for better pneumonia diagnostic aids [9,10]. A target product profile (TPP) was shared with industry, academia, and partners to encourage and guide development of new automated RR counting devices [11]. The ARIDA technical specification listed in the United Nations Children's Fund (UNICEF) request for proposals outlines that any ARIDA must automatically detect and display the RR to aid in the classification of suspected pneumonia in children from the age of 0 to 59 months and include a visual indicator for notification of above or below the age-specific fast breathing thresholds as defined by the WHO IMCI guidelines [3].

In response to the TPP, Philips developed the Children's Respiration Monitor (ChARM) device, which uses an accelerometer-based system to measure the RR in children 0 to 59 months and automatically classifies the breathing rate as fast or normal, based on the age of the child. ChARM is intended to be used by CHWs in low-resource settings. It is strapped around the belly of the child using an elastic belt (Figure 1). ChARM is the first product to be tested as part of the ARIDA field trials, implemented by the Malaria Consortium in Ethiopia and Nepal and sponsored by UNICEF in partnership and with funding from "la Caixa" Banking Foundation.

Figure 1. Illustration of the Philips ChARM device positioned on a child.

In Ethiopia, pneumonia is the number one cause of under-5 mortality, responsible for 16% deaths in children aged younger than 5 years in 2016 [12]. Ethiopia has scaled up iCCM of childhood illness in all regions following a national policy change supporting community-based treatment of childhood pneumonia by CHWs (locally known as health extension workers [HEWs]) in early 2010 [13]. As part of Ethiopia's HEW program, over 42,000 HEWs have been trained for 1 year in iCCM and equipped to assess, classify, and manage pneumonia, malaria, diarrhea, and severe acute malnutrition and provide preventive and curative health services [13]. This paper presents the study design for the evaluation of agreement between an ARIDA and a reference standard RR count for children in Ethiopia.

Methods

Study Aims and Objectives

This study aims to understand whether an ARIDA RR count agrees with an expert panel of pediatricians counting RR by reviewing a video of the child's chest for 60 seconds (reference standard) for children aged younger than 5 years with cough and/or difficult breathing.

The primary objective of this study is to determine the performance of an ARIDA, as defined by agreement and consistency, in children aged younger than 5 years with cough and/or difficulty breathing. The secondary objective is to determine the performance of expert clinicians (ECs) counting RR, as defined by agreement and consistency, in children aged younger than 5 years with cough and/or difficulty breathing. The third objective is to measure RR fluctuation over time after ARIDA device attachment in normal breathing children aged 2 to 59 months.

Study Design

The study is a cross-sectional study comprising three types of RR evaluations: agreement, consistency, and RR fluctuation over time.

Study Site

The study was conducted in the pediatric in- and outpatient departments at Saint Paul's Hospital and Millennium Medical College in Addis Ababa, Ethiopia. This hospital was selected based on the high incidence of pneumonia in outpatient and inpatient departments, interest and willingness of hospital managers to host the study, availability of Integrated Management of Neonatal and Child Illness (IMNCI)-trained [14] expert clinicians (ECs), and availability of a suitable study room, reliable electricity supply, and access to treatment including amoxicillin and oxygen.

Ethics Approval and Consent to Participate

Ethical approval was obtained from the Armauer Hansen Research Institute/ALERT Ethics Review Committee (a biomedical research institute in Ethiopia) on March 7, 2017 (ref. PO02/17) and favorable ethical opinion received by the Liverpool School of Tropical Medicine Ethics Committee. All participants consented to the study by reading and signing the information and consent form.

Study Participants

All children attending in- and outpatient departments at Saint Paul's Hospital and Millennium Medical College in Addis Ababa between April 5 and May 22, 2017, were potential

participants in the study and were systematically screened for eligibility. Children aged 0 to <2 months were excluded from the consistency evaluation due to the anticipated difficulty in attaching two devices at once to a small child. Children aged 0 to <2 months and those with fast breathing were excluded from the fluctuation evaluation due to anticipated difficulty in measuring RR in this group for an extended period of time and also to isolate the effect of the ChARM attachment on RR from other causes of RR fluctuation. All other children aged younger than 5 years who were accompanied by a caregiver aged 18 years and older, not too agitated to be assessed by a research nurse, who did not present with general danger signs or IMNCI referral signs or device manufacturer safety exclusion criteria (wearing supportive device at area of chest/belly, skin not intact in chest/belly, born before 37 weeks of gestation [<2 months only]), were not an inpatient being managed by barrier nursing (such as severe burns, child with neutropenia, severe infectious diseases), and were not advised against research procedures by the supervising clinician were eligible to participate in the study.

General danger signs for newborns (<28 days) include active convulsions or fits, not feeding well, fever (37.5°C [99.5°F] or above), low body temperature (35.5°C [95.9°F] or below), movement only when stimulated, or no movement even when stimulated and for all other age groups include lethargy or unconsciousness, not able to drink/breastfeed, vomiting everything, and active convulsions or fits. IMNCI referral signs for all children include stridor in a calm child, chest indrawing, severe dehydration, severe persistent diarrhea, very severe febrile disease, severe complicated measles, mastoiditis, complicated severe malnutrition, and severe anemia. Written informed consent was obtained from the caregiver before enrollment. Two ECs with extensive experience in assessing and treating children with suspected pneumonia using IMNCI guidelines were selected. They were required to have BSc nursing qualification and an IMNCI certificate.

Outcome Measures

The primary outcome for the agreement evaluation was the mean difference in RR between ARIDA and the reference standard, which summarized the lack of agreement by calculating the average deviation between measures. Similarly for the consistency evaluation, the mean difference in RR between two ARIDA devices was calculate. Table 1 shows all the outcome measures for the study by objective.

- Primary objective: determine the performance of the ARIDA device as defined by agreement and consistency in children aged younger than 5 years with cough or difficulty breathing in a controlled setting
- Secondary objective: determine the performance of ECs counting RR as defined by agreement and consistency in children aged younger than 5 years with cough or difficulty breathing in a controlled setting
- Third objective: measure RR fluctuation over time after ARIDA device attachment in normal breathing children aged 2 to 59 months in a controlled setting

Table 1. Primary objective and outcome measures.

Outcomes	Objective	Evaluation
Primary		
Mean difference between ARIDA[a] and VEP[b] RR[c] count	1	ARIDA agreement
Mean difference between RR counts from two ARIDA devices started simultaneously	1	ARIDA consistency
Secondary		
RMSD[d] between ARIDA and VEP RR counts	1	ARIDA agreement
Percentage ARIDA RR counts ±2 breaths from VEP RR count	1	ARIDA agreement
Absolute mean difference between ARIDA and VEP RR count	1	ARIDA agreement
Positive and negative percentage agreement between ARIDA and VEP RR classification	1	ARIDA agreement
Percentage unsuccessful attempts with ARIDA	1	ARIDA agreement
Percentage failures (3 unsuccessful attempts) with ARIDA	1	ARIDA agreement
Mean time taken to get an ARIDA RR count	1	ARIDA agreement
RMSD between RR counts from two ARIDA devices started simultaneously	1	ARIDA consistency
RMSD of time taken to get ARIDA RR count for two ARIDA devices started simultaneously	1	ARIDA consistency
Mean difference between EC[e] and VEP RR count	2	EC agreement
Mean difference between RR counts from two simultaneous ECs	2	EC consistency
RMSD between RR counts from two simultaneous ECs	2	EC agreement
Percentage EC counts ±2 breaths from VEP RR count	2	EC agreement
Absolute mean difference between EC and VEP RR count	2	EC agreement
Positive and negative percentage agreement between EC and VEP RR classification	2	EC agreement
Percentage unsuccessful attempts for EC RR count	2	EC agreement
Percentage failures (3 unsuccessful attempts) for EC RR count	2	EC agreement
RMSD between RR counts for two simultaneous ECs	2	EC consistency
Difference between RR at baseline and after 1, 3, and 5 minutes after ARIDA attachment	3	Fluctuation
RR trend plotted on a line graph before (baseline) and 1, 3, and 5 minutes after ARIDA attachment	3	Fluctuation

[a]ARIDA: Acute Respiratory Infection Diagnostic Aid (Philips Children's Respiration Monitor device).

[b]VEP: video expert panel (reference standard).

[c]RR: respiratory rate.

[d]RMSD: root mean squared difference.

[e]EC: expert clinician.

Data Collection and Management

Data were collected using an electronic data collection platform (CommCare, Dimagi) installed onto password-protected 7C Pro tablets (Tecno Mobile) and backed up to a protected cloud server. Four-digit unique identification codes were used to anonymize patient data. All videos were transferred using password-protected external hard drives, and paper consent forms were stored in opaque carriers in locked cabinets. All RR evaluation data were entered by two independent research assistants. The data manager downloaded data daily and entered it into a data checker with in-built validation checks. Source videos showing the ARIDA device with the RR count displayed were used to verify ARIDA counts should the two research assistants disagree. Other inconsistencies were rectified by tracing back to paper data entry forms or querying the counts directly with the research team.

Training and Pretest

The video expert panel (VEP) members and ECs were trained for two days on the WHO IMCI method to count RR including practice for half a day using videos of known RR counts [3]. They were individually evaluated using different RR videos to ensure that they were able to count RR ±2 breaths per minute (bpm) from the known RR in 4 out of 5 training videos [15]. All VEP members and ECs passed the competency assessment before starting video review.

Following training, an 8-day pretesting of procedural activities including patient screening, patient flow, and data collection was conducted to ensure the research team was conversant with the data collection procedures, devices, and videography equipment to be used in the study. There was also a pretest of the video panel reference standard evaluations and refresher

training on general danger signs, breath-counting, signs of stridor, and chest indrawing with the IMNCI training video.

Evaluations

Patients were screened by a research nurse in the in- and outpatient departments of the hospital using a screening checklist to ascertain the child's eligibility. An RR classification was made during prescreening by the research nurse using the Mark 2 Acute Respiratory Infection (MK2 ARI) timer to determine whether the child had fast or normal breathing. The prescreening assessment was conducted in a separate part of the hospital and not communicated to the ECs to blind them to the RR classification. Children were enrolled prospectively based on eligibility determined by the screening procedure. The research team then decided, based on the state of the child, age, and breathing status, how many elements of the study to conduct on each child—agreement, consistency, and/or RR fluctuation. Table 2 shows the number of participants aimed to be enrolled to each element of the study by age group and breathing status.

Table 2. Enrollment targets for each type of evaluation by age group and breathing status.

Age group and breathing status (based on screening)	Enrollment targets for three evaluations		
	Agreement	Consistency	Fluctuation
0 to <2 months			
Normal	13	—	—
Fast	39	—	—
2 to <12 months			
Normal	13	13	15
Fast	39	39	—
12 to 59 months			
Normal	13	13	15
Fast	39	39	—

For the agreement evaluation, the research assistant attached an ARIDA to the child and ensured the child was positioned correctly according to device instructions: with his/her back fully supported, either on the lap of the caregiver or lying down on a bed, and the device in line with the child's belly button and one of the nipples [16]. Once the child was calmed, usually by the research assistant clicking their fingers, the videographer started recording and the ARIDA and EC RR count started simultaneously. The EC was blinded to the ARIDA RR count by placing ARIDA on the far side of the child's belly and using a paper cover to shield the screen. The time taken to get an ARIDA count (from when the OK button was pressed to when the device beeped to signal completion of the RR count) was also obtained by a research assistant using a stopwatch. After 60 seconds, if the EC had not obtained an RR count, the EC attempt was recorded as unsuccessful and repeated for both the EC and ARIDA. After 5 minutes or if the ARIDA displayed an error message, the ARIDA attempt was recorded as unsuccessful (with a reason) and repeated for both EC and ARIDA. If the third attempt was still unsuccessful for either device or EC, the evaluation was recorded as a failure. Fifteen different ARIDA devices were used and rotated systematically for all evaluations.

The consistency evaluation followed the same procedure as the agreement evaluation with two ARIDA devices attached to a child using a single belt, positioned in line with each nipple and measured RR from the same starting point. Time taken to obtain each ARIDA RR count was recorded by two research assistants using stopwatches. To measure the consistency between ECs, two ECs conducted separate manual RR counts with MK2 ARI timers over an identical 60-second period. For an EC or ARIDA attempt to be successful, both ECs or both ARIDA devices had to get an RR count. For the RR fluctuation evaluation, an EC counted RR with the MK2 ARI timer for 60 seconds. Following this, the ARIDA was attached to the child and the EC did three more RR counts for 60 seconds in the following time periods: 0 to 1 minute, 2 to 3 minutes, 4 to 5 minutes. On completion of the evaluation, the research team debriefed the caregiver and ensured medical management for the child was completed by the relevant hospital staff.

Reference Standard

The reference standard for the agreement evaluation was a video review by two to four independent VEP members. They were all practicing pediatricians with over 5 years' experience managing pneumonia in children aged younger than 5 years and who had received refresher training in counting RR as per WHO IMCI guidelines (3).

First, two VEP members independently watched a video of the child's chest movements, edited with the layover of the ChARM start and stop sound and a 60 second timer, and counted the number of breaths observed in a full minute. Beep sounds were added by the videographer in sync with the original sounds made by the ChARM device at the start (when the start button on ChARM is pushed) and at the end (when the ChARM displays the result). The sound recorded at the time of recording was also muted to allow the VEP to focus on the sound of the start and stop beeps.

If the first two VEP members agreed (≤±2 bpm), a mean RR count was used as the reference standard. If they disagreed (>±2 bpm), a third VEP member reviewed the video and if two out of three counts agreed (≤±2 bpm), the mean of the two closest RR counts was used. If all three VEP members disagreed (>±2

bpm), the video was sent for review to a fourth VEP member. If the fourth VEP member's count agreed (≤±2 bpm) with any of the first three VEP members' counts, the mean of the two closest counts was used. If all four panel members disagreed (>±2 bpm), the data from this evaluation were excluded from the agreement analysis.

Sample Size

The primary outcome on which sample size was based was the agreement between the ARIDA and VEP RR counts. As per Bland-Altman [17], we conducted a precision-based sample size calculation based on the confidence interval for the 95% limits of agreement. The formula estimates the required number of children per age group (n) based on the desired width of the confidence interval. Using normal approximation and allowing a confidence interval of ±0.5 standard deviations of the difference between the two devices, a sample size of 46 children per age group was required for the agreement and consistency evaluations, adjusted to 52 per group for failure to get a reference standard count. For the RR fluctuation evaluation, a sample size of 30 children was used.

Data Analysis

Data analysis for all three RR evaluations was conducted in Stata 13 (StataCorp LLC) and Excel (Microsoft Corp). First, the number of children screened, eligible, consented, and enrolled in each type of evaluation was described. Baseline characteristics (age and sex) by screening breathing status (normal/fast) for those enrolled were described. All full-length source videos were reviewed for quality assurance purposes, and descriptive information on the video quality was recorded, including those where all four VEP members disagreed on the RR count. For the ARIDA and EC agreement and consistency evaluations, mean difference, root mean square difference, absolute mean difference, proportion of RR counts ±2 bpm from the reference standard, and positive and negative percentage agreement with 95% confidence intervals were calculated in Stata 13 by age group, and Bland-Altman plots with limits of agreement and 95% confidence intervals by age group and breathing status were created. Percentage of unsuccessful attempts and failures (defined by three unsuccessful attempts) and mean time to get an ARIDA RR count were calculated. A per-protocol analysis was used whereby children were excluded from the analysis if an RR could not be obtained simultaneously by the ARIDA and by the EC, with a VEP RR reading where at least two of the panel members were within ±2 bpm of each other. For the RR fluctuation evaluation, mean difference in the RR count between baseline and 1 minute, 1 and 3 minutes, and 3 and 5 minutes were calculated. The proportion of children with fast or normal RR classification at baseline and the change between RR classifications over time were analyzed.

Quality Assurance, Supervision, and Monitoring

Malaria Consortium and UNICEF (Supply Division and Ethiopia Country Office) conducted quality assurance visits every 2 weeks to the research site during data collection. All data collected from the screening and RR evaluations were checked and verified by the data manager daily. A sample of three videos was sent weekly to an independent study advisor for RR evaluation using WHO IMCI guidelines [3] and to Malaria Consortium HQ for quality assurance. The project had an 11-person Advisory Committee made up of experts on maternal and child health who provided technical oversight and reviewed the study protocol.

Results

The project was funded in 2016. Data were collected between 5 April until 22 May 2017. Authors are drafting the results for publication.

Discussion

Accurately diagnosing pneumonia in children aged younger than 5 years remains a significant problem in resource-poor settings. Manually counting RR is inherently challenging for CHWs, resulting in both over and under diagnosis and treatment. This diagnostic performance study in Ethiopia aims to provide evidence for the performance of an ARIDA device when used in a controlled setting.

Evaluating performance of new RR counters is difficult due to the absence of an appropriate gold standard. Selecting a robust reference standard when designing this study was a challenge. The aim was to have one reference standard for evaluating any new ARIDA regardless of the technology for calculating RR. Retrospective review of video recordings by a panel of experts has been used as a reference standard for a number of pneumonia studies [18-20]. It allows many experts to assess the same patient, thus limiting bias that could arise from having one expert per child and reducing the number of experts present in the room, whose presence could agitate the child and affect their RR. It also allows the expert to review the evaluation numerous times and adjust the speed and zoom of the video to aid the counter. An interrater agreement study in northeast Tanzania measured the agreement between two pediatricians reviewing RR videos of children aged 2 to 59 months. They found that in two-thirds of cases, both pediatricians agreed on the RR within ±2bpm, which represents fair agreement (kappa=.34) and in ninety-six percent of cases, both pediatricians agreed on RR classification, representing perfect agreement (kappa=.85) [21]. Recognizing the limitations of humans counting RR using a video, in the absence of a gold standard and with recommendations from the Advisory Committee, the video reference standard was selected.

A strength of this study is that the design allows for contemporaneous comparison between the RR count from the ARIDA, EC, and video reference standard. While the comparison is imperfect, as the ARIDA can take longer than 60 seconds to obtain a count compared with the VEPs and ECs who assessed RR over 60 seconds, it remains useful for identifying increased RR and therefore whether the RR classification of breathing status is comparable.

To minimize RR counting errors, this study was implemented with two days of training and an assessment for the VEP members and ECs to ensure a consistent methodology for RR counting. Interrater agreement between humans could be improved with guidance about how to define a breath versus a

movement and additional standardization between humans through training and practice using this guidance to count RR for a selection of videos. Furthermore, a video annotation aid that allows the panel member to mark breaths and non–breath movements directly on the video could reduce human RR counting inconsistencies and allow for discussion and consensus building between panel members about the RR of videos.

Mean difference with 95% confidence intervals was selected as the agreement measure for the primary outcome. A disadvantage of this measure is that positive and negative bias cancel out to give a lower mean difference. An alternative is to use the limits of agreement with 95% confidence intervals on the Bland-Altman plot to visually show the agreement between the two measures and estimate the precision of the estimates. We recommend that global consensus and guidance on an acceptable level of agreement between a new automated RR counter and a reference standard as measured by the range of the limits of agreement on a Bland-Altman plot is sought in addition to global consensus on the reference standard methodology. This will provide ministries of health with evidence to make an informed decision on the performance of new RR devices to inform introduction and scale up of these devices.

Acknowledgments

We would like to thank St Paul's Hospital Millennium Medical College for allowing us to conduct the research at their hospital and allowing their staff to be available for the VEP. We also extend thanks to the videographer, Tewodros Emiru, Dr Hollie Spence who helped in the additional review of video references, and the research team: Yenealem Reta, Fiseha Mihretu, Ruth Haiu, Okulu Sultan, Hanna Birkneh, Lidiya Zerihun, Atresaw Zegye, Kalkidan Belay, Sileshi Tsegaye, and Aynalem Hailem. We are also grateful for the continued support from the Ethiopian Federal Ministry of Health and members of the Advisory Committee: Quique Bassat, David Milestone, Kristoffer Gandrup-Marino, Gian Gandhi, Mark Young, Leith Greenslade, Shamin Qazi, Jim Black, Raza Izadnegahdar, Salim Sadruddin, and HRH the Infanta Cristina of Spain. The study was sponsored by UNICEF in partnership and with funding from "la Caixa" Banking Foundation. The manufacturer of the device (Philips) did not contribute any funding for this work and did not participate in any data analysis or interpretation. The Malaria Consortium research team has subsequently received funding from the Philips Foundation to conduct future work on developing an improved respiratory rate reference standard.

Authors' Contributions

KK, KB, CW, SM, DG, JHB, PL, and NM designed the study protocol. CW, KB, SM, and KK created the data collection tools and standard operating procedures with review from TH, CM, and JHB. TH, CW, KB, SM, and KK implemented the data collection. MP, KK, CW, KB, and SM designed and implemented the data analysis. CW drafted the paper. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. McAllister D, Liu L, Shi T, Chu Y, Reed C, Burrows J, et al. Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. Lancet Glob Health 2019 Jan;7(1):e47-e57 [FREE Full text] [doi: 10.1016/S2214-109X(18)30408-X] [Medline: 30497986]
2. Requejo J, Bryce J, Barros A, Berman P, Bhutta Z, Chopra M, et al. Countdown to 2015 and beyond: fulfilling the health agenda for women and children. Lancet 2015 Jan 31;385(9966):466-476. [doi: 10.1016/S0140-6736(14)60925-9] [Medline: 24990815]
3. World Health Organization. Integrated management of childhood illness: chart booklet 2014 URL: http://apps.who.int/iris/bitstream/10665/104772/16/9789241506823_Chartbook_eng.pdf [accessed 2020-02-10]
4. Young M, Wolfheim C, Marsh D, Hammamy D. World Health Organization/United Nations Children's Fund joint statement on integrated community case management: an equity-focused strategy to improve access to essential treatment services for children. Am J Trop Med Hyg 2012 Nov;87(5 Suppl):6-10 [FREE Full text] [doi: 10.4269/ajtmh.2012.12-0221] [Medline: 23136272]
5. Mukanga D, Babirye R, Peterson S, Pariyo G, Ojiambo G, Tibenderana J, et al. Can lay community health workers be trained to use diagnostics to distinguish and treat malaria and pneumonia in children? Lessons from rural Uganda. Trop Med Int Health 2011 Oct;16(10):1234-1242 [FREE Full text] [doi: 10.1111/j.1365-3156.2011.02831.x] [Medline: 21752163]
6. Källander K, Tomson G, Nsabagasani X, Sabiiti JN, Pariyo G, Peterson S. Can community health workers and caretakers recognise pneumonia in children? Experiences from western Uganda. Trans R Soc Trop Med Hyg 2006 Oct;100(10):956-963. [doi: 10.1016/j.trstmh.2005.11.004] [Medline: 16455119]
7. Muro F, Mtove G, Mosha N, Wangai H, Harrison N, Hildenwall H, et al. Effect of context on respiratory rate measurement in identifying non-severe pneumonia in African children. Trop Med Int Health 2015 Jun;20(6):757-765 [FREE Full text] [doi: 10.1111/tmi.12492] [Medline: 25728867]

8. UNICEF. ARIDA (Acute Respiratory Infection Diagnostic Aid) 2017 URL: https://www.unicef.org/innovation/innovation_81722.html [accessed 2020-02-10]
9. Ginsburg AS, Sadruddin S, Klugman KP. Innovations in pneumonia diagnosis and treatment: a call to action on World Pneumonia Day, 2013. Lancet Glob Health 2013 Dec;1(6):e326-e327 [FREE Full text] [doi: 10.1016/S2214-109X(13)70117-7] [Medline: 25104591]
10. Källander K, Young M, Qazi S. Universal access to pneumonia prevention and care: a call for action. Lancet Respir Med 2014 Dec;2(12):950-952. [doi: 10.1016/S2213-2600(14)70248-6] [Medline: 25466341]
11. UNICEF. Target Product Profile: acute respiratory infection diagnostic aid (ARIDA). Copenhagen: UNICEF Supply Division; 2014 Nov 14. URL: https://www.unicef.org/videoaudio/PDFs/ARIDA_-_Target_Product_Profile_(2).pdf [accessed 2020-02-11]
12. UNICEF. Estimates of child cause of death, acute respiratory infection 2018 URL: https://data.unicef.org/wp-content/uploads/2018/02/CoD_ARI_Feb-2018_WHO_MCEE_236.xlsx [accessed 2020-02-11]
13. Legesse H. National scale-up of integrated community case management in rural Ethiopia: implementation and early lessons learned. Ethiop Med J 2014 Oct;52 Suppl 3:15-26. [Medline: 25845070]
14. World Health Organization. 2014 Mar. Integrated management of newborn and childhood illness chart booklet URL: https://tinyurl.com/wnhee4m [accessed 2020-03-02]
15. USAID. 2015. Project CLEAR: pneumonia pretesting outputs URL: https://tinyurl.com/uf8xv55 [accessed 2020-03-02]
16. Philips. ChARM Instructions for use 2016 URL: https://tinyurl.com/t27txn4 [accessed 2020-02-11]
17. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 08;1(8476):307-310. [Medline: 2868172]
18. Sinyangwe C, Graham K, Nicholas S, King R, Mukupa S, Källander K, et al. Assessing the quality of care for pneumonia in integrated community case management: a cross-sectional mixed methods study. PLoS One 2016;11(3):e0152204 [FREE Full text] [doi: 10.1371/journal.pone.0152204] [Medline: 27011331]
19. Black J, Gerdtz M, Nicholson P, Crellin D, Browning L, Simpson J, et al. Can simple mobile phone applications provide reliable counts of respiratory rates in sick infants and children? An initial evaluation of three new applications. Int J Nurs Stud 2015 May;52(5):963-969. [doi: 10.1016/j.ijnurstu.2015.01.016] [Medline: 25712876]
20. Karlen W, Gan H, Chiu M, Dunsmuir D, Zhou G, Dumont GA, et al. Improving the accuracy and efficiency of respiratory rate measurements in children using mobile devices. PLoS One 2014;9(6):e99266 [FREE Full text] [doi: 10.1371/journal.pone.0099266] [Medline: 24919062]
21. Muro F, Mosha N, Hildenwall H, Mtei F, Harrison N, Schellenberg D, et al. Variability of respiratory rate measurements in children suspected with non-severe pneumonia in north-east Tanzania. Trop Med Int Health 2017 Feb;22(2):139-147 [FREE Full text] [doi: 10.1111/tmi.12814] [Medline: 27862739]

Abbreviations

ARI: acute respiratory infection
ARIDA: Acute Respiratory Infection Diagnostic Aid
bpm: breaths per minute
ChARM: Children's Respiration Monitor
CHW: community health worker
EC: expert clinician
FLHFW: first-level health facility worker
HEW: health extension worker
iCCM: Integrated Community Case Management
IMCI: Integrated Management of Childhood Illness
IMNCI: Integrated Management of Neonatal and Child Illness
RR: respiratory rate
TPP: target product profile
UNICEF: United Nation Children's Fund
VEP: video expert panel
WHO: World Health Organization

Edited by G Eysenbach; submitted 16.10.19; peer-reviewed by ECY Su, S Rhedin, S Sanchez; comments to author 30.11.19; revised version received 03.01.20; accepted 14.01.20; published 02.04.20.

Please cite as:
Ward C, Baker K, Marks S, Getachew D, Habte T, McWhorter C, Labarre P, Howard-Brand J, Miller NP, Tarekegn H, Deribessa SJ, Petzold M, Kallander K
Determining the Agreement Between an Automated Respiratory Rate Counter and a Reference Standard for Detecting Symptoms of Pneumonia in Children: Protocol for a Cross-Sectional Study in Ethiopia
JMIR Res Protoc 2020;9(4):e16531
URL: https://www.researchprotocols.org/2020/4/e16531
doi:10.2196/16531
PMID:32238340

Protocol

Effectiveness of Gastric Bypass Versus Gastric Sleeve for Cardiovascular Disease: Protocol and Baseline Results for a Comparative Effectiveness Study

Karen J Coleman[1], BS, PhD; Heidi Fischer[1], PhD; David E Arterburn[2], MD, MPH; Douglas Barthold[3], PhD; Lee J Barton[1], MS; Anirban Basu[3], PhD; Anita Courcoulas[4], MD; Cecelia L Crawford[5], DNP, RN; Peter Fedorka[6], MD; Benjamin Kim[7], MD; Edward Mun[7], MD; Sameer Murali[8], MD; Kristi Reynolds[1], PhD; Kangho Suh[3], PharmD; Rong Wei[1], MS; Tae K Yoon[1], MS; Robert Zane[7], MD

[1]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA, United States

[2]Health Research Institute, Kaiser Permanente Washington, Seattle, WA, United States

[3]The Comparative Health Outcomes, Policy, and Economics Institute, Department of Pharmacy, University of Washington, Seattle, WA, United States

[4]Department of Surgery, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[5]Regional Nursing Research Program, Kaiser Permanente Southern California, Pasadena, CA, United States

[6]Department of Surgery, San Bernardino Medical Center, Kaiser Permanente Southern California, Ontario, CA, United States

[7]Department of Surgery, South Bay Medical Center, Kaiser Permanente Southern California, Harbor City, CA, United States

[8]Center for Healthy Living, San Bernardino Medical Center, Kaiser Permanente Southern California, Fontana, CA, United States

Corresponding Author:
Karen J Coleman, BS, PhD
Department of Research and Evaluation
Kaiser Permanente Southern California
100 S Los Robles
Pasadena, CA
United States
Phone: 1 6265643580
Email: Karen.J.Coleman@kp.org

Abstract

Background: When compared with conventional weight loss strategies, bariatric surgery results in substantially greater durable weight loss and rates of disease remission.

Objective: The ENGAGE CVD (Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease) cohort study aimed to provide population-based, comprehensive, rigorous evidence for clinical and policy decision making regarding the choice between gastric bypass and gastric sleeve for overall cardiovascular disease (CVD) risk reduction, risk factor remission, and safety.

Methods: The cohort had 22,095 weight loss surgery patients from a large integrated health care system in Southern California assembled from 2009 to 2016 who were followed up through 2018. Bariatric surgery patients were followed up for the length of their membership in the health care system. Of the patients who had at least five years of follow-up (surgery between 2009 and 2013), 85.86% (13,774/16,043) could contribute to the outcome analyses for the ENGAGE CVD cohort.

Results: Patients in the ENGAGE CVD cohort were 44.6 (SD 11.4) years old, mostly women (17,718/22,095; 80.19%), with 18.94% (4185/22,095) non-Hispanic black and 41.80% (9235/22,095) Hispanic, and had an average BMI of 44.3 (SD 6.9) kg/m^2 at the time of surgery. When compared with patients who did not contribute data to the 5-year outcome analysis for the ENGAGE CVD cohort (2269/16,043; 14.14%), patients who contributed data (13,774/16,043; 85.86%) were older (P=.002), more likely to be women (P=.02), more likely to be non-Hispanic white (P<.001), more likely to have had an emergency department visit in the year before surgery (P=.006), less likely to have a mental illness before surgery (P<.001), and more likely to have had a CVD event at any time before surgery (P<.001).

Conclusions: This study had one of the largest populations of gastric sleeve patients (n=13,459). The 5-year follow-up for those patients who had surgery between 2009 and 2013 was excellent for a retrospective cohort study at 85.86% (13,774/16,043).

Unlike almost any study in the literature, the majority of the ENGAGE CVD cohort was racial and ethnic minority, providing a rare opportunity to study the effects of bariatric surgery for different racial and ethnic groups, some of whom have the highest rates of severe obesity in the United States. Finally, it also used state-of-the-art statistical and econometric comparative effectiveness methods to mimic the effect of random assignment and control for sources of confounding inherent in large observational studies.

International Registered Report Identifier (IRRID): RR1-10.2196/14936

(JMIR Res Protoc 2020;9(4):e14936) doi:10.2196/14936

KEYWORDS

race; weight loss surgery; integrated health care system

Introduction

Overview of Surgical Treatment for Severe Obesity

The prevalence of severe obesity (BMI >35 kg/m^2) has increased over the past several decades. Rates are as high as 36% for middle-aged black women compared with 16% for their white counterparts in the United States [1]. Even with intensive, multicomponent lifestyle interventions, only 50% of studies show 5% weight loss (considered clinically meaningful), and most of the participants gain back at least half of this lost weight over 18 to 30 months [2]. These poor outcomes have resulted in the development of surgical treatments, referred to as bariatric surgery, for severe obesity. When compared with conventional weight loss strategies, bariatric surgery results in seven times the amount of weight loss and 15.8 times the rate of diabetes remission [3], and these differences remain up to 5 years [4,5]. Given the poor results from traditional weight loss methods [2], and the designation of obesity as a disease [6], bariatric surgery may become a more common treatment of choice for adults with severe obesity.

Two surgical treatments constitute most bariatric operations in the United States: vertical sleeve gastrectomy (VSG) and Roux-en-Y gastric bypass (RYGB). VSG, in which stomach size is reduced, was initially performed as the first part of a multistage procedure in 2000. RYGB, in which gastric capacity is also limited but with an additional bypass of the first few feet of small intestine, was first performed in 1994 [7]. VSG has emerged as the fastest growing bariatric operation in the United States. Between 2008 and 2014, there was a dramatic increase in VSGs from 4% to 51% of all bariatric operations, whereas RYGB declined from 51% to 27% [8]. The reasons for this shift have not been systematically studied, but based upon our own work [9], it is likely because of patients' and surgeons' perceptions that although VSG and RYGB have similar weight loss and disease remission, VSG is easier to perform with fewer complications compared with RYGB.

Evidence for Comparative Effectiveness of Surgical Treatments

Unfortunately, the use of VSG has outpaced a rigorous evidence base for its comparative effectiveness to RYGB [10-15]. In addition, few large population-based studies in real-world health care settings have adequate methodological rigor to account for the fact that VSG and RYGB operations are not randomly assigned. Patients with risk factors for cardiovascular disease (CVD), especially type 2 diabetes mellitus (T2DM), are more likely to undergo RYGB [16]. The reasons for this are not clear; however, it is likely that surgeons and patients believe RYGB is more effective than VSG for resolving T2DM. If this treatment choice preference is not accounted for in the analyses, then erroneous conclusions could be made about the effectiveness of one operation compared with another because the patients receiving each treatment are different in ways that also affect the outcome.

Addressing Limitations in the Evidence Base

Rigorous statistical methods such as matching, propensity scores, and/or instrumental variables have only been applied to the study of the comparative effectiveness of VSG and RYGB in the remission and relapse of T2DM. To our knowledge, there have been no rigorous comparative effectiveness studies published for other risk factors for CVD, including hypertension and dyslipidemia. In addition, there are no published studies on the comparative effectiveness of VSG and RYGB for reducing overall CVD risk beyond the first year after surgery. The ENGAGE CVD (Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease) cohort study was funded by the National Heart, Lung, and Blood Institute to provide population-based, comprehensive, rigorous evidence for clinical and policy decision making regarding the choice between RYGB and VSG for overall CVD risk reduction, risk factor remission, and safety. The ENGAGE CVD study uses state-of-the-art statistical and econometric comparative effectiveness methods, including propensity scores and local instrumental variables (LIVs), to mimic the effect of random assignment and control for sources of both observed and unobserved confounding inherent in large observational studies.

Study Objectives and Hypotheses

There were three aims for the ENGAGE CVD study. Aim 1 compared the effectiveness of VSG and RYGB in remission and relapse of CVD risk factors and reduction in overall CVD risk. For this aim we hypothesized that RYGB patients would experience a higher rate of T2DM, hypertension, and dyslipidemia remission and lower rate of relapse compared with VSG patients. RYGB patients would also have a greater reduction in overall CVD risk. Aim 2 compared VSG and RYGB surgical safety. We hypothesized that VSG patients would have better short- and long-term safety outcomes than RYGB patients. Aim 3 was designed to understand the treatment effect heterogeneity in remission and relapse of CVD risk factors, reduction in overall CVD risk, and safety outcomes for patients with different racial and ethnic backgrounds, genders, ages, and disease burdens at the time of surgery. Based upon our own work in this area, we expected an interaction of racial and ethnic

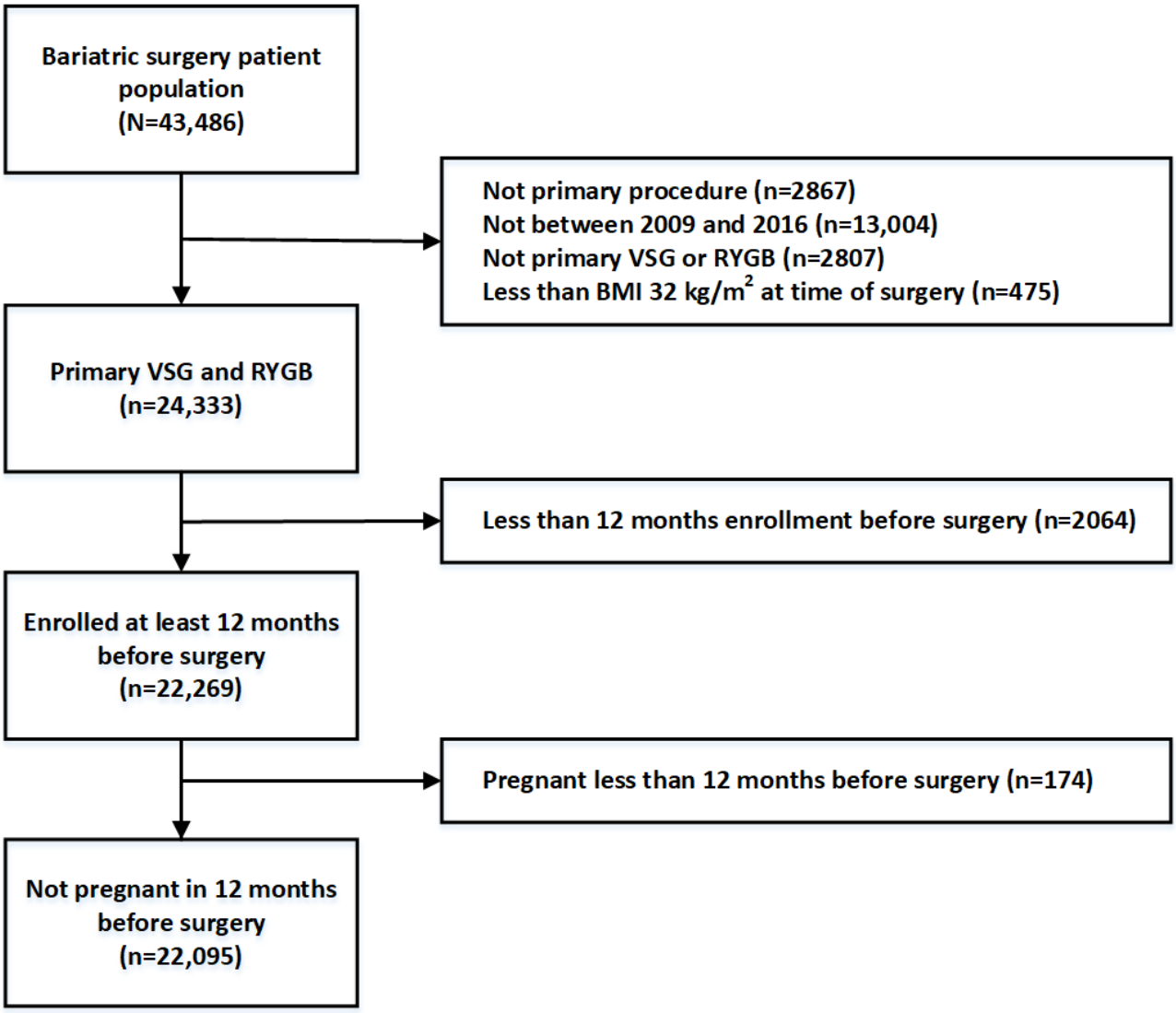
minority, male sex, older age, and having a higher disease burden in attenuating the differences hypothesized between RYGB and VSG.

Methods

Settings and Participants

Figure 1 shows the process of selecting the ENGAGE CVD cohort (n=22,095) and Table 1 presents descriptive statistics for the RYGB and VSG patients in the cohort. The cohort was assembled from 2009 to 2016 from a large integrated health care system serving the Southern California region of the United States. This health care system had 4.2 million members, 14 hospitals, 200 medical offices, 5700 physicians, and 23 bariatric surgeons at the time the cohort was assembled. Inclusion and exclusion criteria for the ENGAGE CVD cohort are shown in Figure 1.

Figure 1. Inclusion and exclusion criteria for the ENGAGE CVD (Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease) cohort study. RYGB: Roux-en-Y gastric bypass; VSG: vertical sleeve gastrectomy.



This cohort of bariatric surgery patients was similar to other bariatric studies published in the United States, with the exception that there was a much higher proportion of ethnic/racial minorities (63.8%) than in other published work [17,18]. Eligibility for weight loss surgery in this health care system was based upon national recommendations [19]: Having a BMI ≥40 kg/m^2 or having a BMI of 35-39 kg/m^2 and at least one obesity-related comorbid condition such as sleep apnea, T2DM, and heart disease. Patients meeting these criteria could still be refused surgery if the surgeon determined that the patient had excessively high medical risk for surgery and in some cases, patients could have surgery if their BMI was as low as 32 kg/m^2 with T2DM. Only 3.66% (808/22,095) of the ENGAGE CVD cohort had a BMI of 32-34.99 kg/m^2 at the time of their operation.

Measures

Bariatric surgery patients were followed up for the length of their membership in the health care system. Weight, height, and blood pressure were measured at every outpatient visit. In general, laboratory measures relevant to CVD such as glucose and glycated hemoglobin (HbA_{1c}) were measured before surgery and at least annually following surgery. Lipids were only measured routinely every 5 years following national screening

guidelines [20]. All data were abstracted from the electronic health record for the period of 2009 to 2018 and comprised the following broad categories of information.

Baseline

At the time of surgery, patient self-reported date of birth, gender, and race/ethnicity were obtained. Details of surgery type, surgeon, and surgery location were also assembled for the cohort.

Baseline and Follow-Up

Data were obtained for 24 months before the date of surgery and up to 10 years after surgery and included the following: (1) Dates and status of health care system enrollment and types of insurance coverage including pharmacy coverage; (2) vital signs such as height, weight, and blood pressure (in general, height was self-reported, and weight and blood pressure were measured by clinical staff at every outpatient visit. Previous research in health care settings has demonstrated that heights and weights from electronic medical records are valid and suitable for research [21]. Most blood pressure measurements were performed by certified medical assistants using automatic devices.); (3) self-reported smoking status from outpatient visits; (4) comprehensive prescription data for each drug dispensed at health care system pharmacies and all outpatient and inpatient laboratory results were also available (almost all patients [>96%] had benefits that incentivized the use of health care system pharmacies and laboratories); (5) all dates and types of health care utilization for inpatient, emergency department, and outpatient settings (including external claims data from contracted surgical providers); and (6) the diagnoses and procedures associated with this health care utilization.

Outcomes

The primary outcome for the first aim of the ENGAGE CVD study was T2DM remission and relapse in each of the years of follow-up after bariatric surgery up to 5 years. Secondary outcomes for aim 1 were hypertension and dyslipidemia remission and relapse, and overall 10-year CVD risk as assessed with the new American College of Cardiology and the American Heart Association guidelines, referred to as the Pooled Cohort Equations Risk Calculator or ASCVD risk score [22], over this same time period. The primary outcome for aim 2 was a 30-day composite measure of major adverse events specific to bariatric surgery patients. The secondary outcomes for aim 2 were long-term annual rates of reoperations/revisions, readmissions, emergency department use, and all-cause mortality up to 5 years following bariatric surgery.

Analyses

Summary statistics for the ENGAGE CVD cohort were generated using means and standard deviations for continuous variables and frequency and percent for categorical variables. Unadjusted differences between patients who were alive and still members of the health care system 5 years after surgery (n=13,774) and those patients who were not (n=2269), as well as between patients who had RYGB (n=8636) and VSG (n=13,459) were analyzed with independent sample t tests (continuous); and the Chi-square statistic and Kruskal-Wallis test (categorical).

The main analysis for the outcomes was a LIV approach [23]. This approach used a continuous instrumental variable to estimate the effect on every margin of the patient population and estimated population average effects to understand how different patients did with different treatments. This is referred to as heterogeneity of treatment effects (HTE) [24]. A clinically intuitive description of these methods applied to a clinical setting has been recently published [25]. These findings were compared with more traditional comparative effectiveness methods in retrospective observational studies, such as inverse-probability weighted propensity score regression [26], that only controlled for observed confounders in the decision between VSG and RYGB operations.

Results

Participants

Descriptive characteristics for patients in the ENGAGE CVD cohort are shown in Table 1. Overall, the cohort was 44.6 (SD 11.35) years old, with 80.19% (17,718/22,095) women, 18.94% (4185/22,095) non-Hispanic blacks, and 41.80% (9235/22,095) Hispanics. Patients had an average BMI of 44.30 (SD 6.88) kg/m^2 at surgery with the majority having a BMI between 35-50 kg/m^2 (17,386/22,095; 78.69%). In the 2 years before surgery, patients had been diagnosed with the following conditions: 36.56% (8078/22,095) gastroesophageal reflux disease (GERD), 15.48% (3421/22,095) sleep apnea, 36.72% (8114/22,095) T2DM, 53.80% (11,887/22,095) hypertension, 72.82% (16,090/22,095; dyslipidemia, and 11,967/22,095; 54.16% mental health condition (primarily depression). Only 3.44% (759/22,095) had a CVD event in their lifetime before surgery.

Missing Data

Of the 22,095 patients in the ENGAGE CVD cohort, 16,043 (72.61%) had surgery between 2009 and 2013 and thus had enough follow-up time for the assessment of outcomes at 5 or more years following surgery. Of these 16,043 patients, 13,774 (85.86%) were still living (104 died before 5 years) and members of the health plan (2165 discontinued membership 5 years after surgery) at 5 years after surgery. Table 2 presents differences in baseline data for the ENGAGE CVD cohort of patients who were alive and still members of the health care system 5 years after surgery (n=13,774) compared with those patients who were not (n=2269). When compared with patients who did not contribute data to the 5-year outcome analysis for the ENGAGE CVD cohort (2269/16,043; 14.14%), patients who contributed data (13,774/16,043; 85.86%) were older (P=.002), more likely to be women (P=.02), more likely to be non-Hispanic white (P<.001), more likely to have a duodenal ulcer at the time of surgery (P<.001), less likely to have dyslipidemia (P<.001), more likely to have had an emergency department visit in the year before surgery (P=.006), less likely to have a mental illness before surgery (P<.001), and more likely to have had a CVD event at any time before surgery (P<.001).

Table 1. Characteristics of patients before surgery who are included in the ENGAGE CVD (Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease) cohort study (n=22,095) by bariatric operation (vertical sleeve gastrectomy [VSG] and Roux-en-Y gastric bypass [RYGB]. Characteristics at the time of surgery are compared between VSG and RYGB.

Characteristics	Overall (n=22,095)	VSG[a] (n=13,459)	RYGB[b] (n=8636)	*P* value
Age years), mean (SD)	44.6 (11.35)	44.1 (11.39)	45.4 (11.25)	<.001
Women, n (%)	17,718 (80.19)	10,850 (80.62)	6868 (79.53)	.06
Non-Hispanic black, n (%)	4185 (18.94)	2886 (21.44)	1299 (15.04)	<.001
Hispanic, n (%)	9235 (41.80)	5576 (41.43)	3659 (42.37)	<.001
Non-Hispanic white, n (%)	7997 (36.19)	4599 (34.17)	3398 (39.35)	<.001
Other, n (%)	678 (3.07)	398 (2.96)	280 (3.24)	<.001
Weight loss in the year before surgery (lbs), mean (SD)	−17.2 (14.74)	−17.8 (14.53)	−16.4 (15.02)	<.001
BMI at surgery (kg/m^2), mean (SD)	44.3 (6.88)	43.8 (6.63)	45.1 (7.17)	<.001
BMI 32-34.99 kg/m^2 at surgery, n (%)	808 (3.66)	547 (4.06)	261 (3.02)	<.001
BMI 35-39.99 kg/m^2 at surgery, n (%)	5531 (25.03)	3633 (26.99)	1898 (21.98)	<.001
BMI 40-49.99 kg/m^2 at surgery, n (%)	11,856 (53.66)	7185 (53.38)	4671 (54.09)	<.001
BMI >50 kg/m^2 at surgery, n (%)	3900 (17.65)	2094 (15.56)	1806 (20.91)	<.001
Any lifetime cardiovascular disease event before surgery, n (%)	759 (3.44)	421 (3.13)	338 (3.91)	.002
Gastroesophageal reflux disease in 2 years before surgery, n (%)	8078 (36.56)	4472 (33.23)	3606 (41.76)	<.001
Esophagitis in 2 years before surgery, n (%)	388 (1.76)	217 (1.61)	171 (1.98)	.04
Gastric ulcer in 2 years before surgery, n (%)	153 (0.07)	96 (0.071)	57 (0.07)	.64
Duodenal ulcer in 2 years before surgery, n (%)	1411 (6.39)	816 (6.06)	595 (6.89)	.01
Peptic ulcer in 2 years before surgery, n (%)	346 (1.57)	197 (1.46)	149 (1.73)	.13
Gastritis duodenitis in 2 years before surgery, n (%)	2538 (11.49)	1518 (11.28)	1020 (11.81)	.23
Dyspepsia in 2 years before surgery, n (%)	2625 (11.88)	1614 (11.99)	1011 (11.71)	.52
Hiatal hernia in 2 years before surgery, n (%)	688 (3.11)	382 (2.84)	306 (3.54)	.003
Gastrointestinal bleed in 2 years before surgery, n (%)	9 (0.00)	5 (0.00)	4 (0.00)	.74
Aspirin use in 1 year before surgery, n (%)	3925 (17.76)	1875 (13.93)	2050 (23.74)	<.001
Aspirin use in 3 months before surgery, n (%)	2517 (11.39)	1255 (9.32)	1262 (14.61)	<.001
NSAID[c] use in 1 year before surgery, n (%)	9630 (43.58)	5916 (43.96)	3714 (43.01)	.17
NSAID use in 3 months before surgery, n (%)	3260 (14.75)	1985 (14.75)	1275 (14.76)	.975
Cirrhosis in 2 years before surgery, n (%)	122 (0.01)	77 (0.01)	45 (0.01)	.62
Sleep apnea in 2 years before surgery, n (%)	3421 (15.48)	1983 (14.73)	1438 (16.65)	<.001
Type 2 diabetes mellitus in 2 years before surgery, n (%)	8114 (36.72)	3827 (28.43)	4287 (49.64)	<.001
Hypertension in 2 years before surgery, n (%)	11,887 (53.80)	6704 (49.81)	5183 (60.01)	<.001
Chronic kidney disease in 2 years before surgery, n (%)	2623 (11.87)	1402 (10.42)	1221 (14.14)	<.001
Dyslipidemia in 2 years before surgery, n (%)	16,090 (72.82)	9409 (69.90)	6681 (77.36)	<.001
Any mental health condition in 2 years before surgery, n (%)	11,967 (54.16)	7153 (53.15)	4814 (55.74)	<.001
Attendance rate in 1 year before surgery (range 0%-100%), mean (SD)	76.60 (12.39)	76.30 (12.31)	77.00 (12.49)	<.001
Any inpatient visit 1 year before surgery, n (%)	1317 (5.96)	722 (5.36)	595 (6.89)	<.001
Any emergency department visit in 1 year before surgery, n (%)	4655 (21.07)	2788 (20.71)	1867 (21.62)	.11

[a]VSG: vertical sleeve gastrectomy.

[b]RYGB: Roux-en-Y gastric bypass.

[c]NSAID: nonsteroidal anti-inflammatory drug.

Table 2. Characteristics of patients before surgery in the ENGAGE CVD (Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease) cohort study who accumulated 5 years of follow-up after surgery (n=16,043). Findings are compared for those patients who had missing (2269/16,043; 14.14%) and no missing (13,774/16,043; 85.86%) data at 5 years following bariatric surgery.

Variables	Accumulated 5 years of follow-up (N=16,043)	Missing 5-year data (N=2269)	Complete 5-year data (N=13,774)	*P* value
Roux-en-Y gastric bypass, n (%)	6891 (42.95)	1104 (48.66)	5787 (42.01)	<.001
Vertical sleeve gastrectomy, n (%)	9152 (57.05)	1165 (51.34)	7987 (57.99)	<.001
Age (years), mean (SD)	44.1 (11.92)	45.0 (9.54)	44.0 (12.26)	.002
Women, n (%)	12,860 (80.16)	1779 (78.40)	11,081 (80.45)	.02
Non-Hispanic black, n (%)	3067 (19.12)	375 (16.52)	2692 (19.54)	<.001
Hispanic, n (%)	6470 (40.33)	922 (40.64)	5548 (40.28)	<.001
Non-Hispanic white, n (%)	6011 (37.47)	929 (40.94)	5082 (36.90)	<.001
Other, n (%)	495 (3.09)	43 (1.90)	452 (3.28)	<.001
Weight loss in year before surgery (lbs), mean (SD)	−16.9 (15.01)	−17.3 (14.94)	−16.9 (15.02)	.04
BMI at surgery (kg/m^2), mean (SD)	44.7 (6.95)	44.8 (7.17)	44.6 (6.91)	.59
BMI 32-34.99 kg/m^2 at surgery, n (%)	508 (3.17)	69 (3.04)	439 (3.19)	.84
BMI 35-39.99 kg/m^2 at surgery, n (%)	3701 (23.07)	516 (22.74)	3185 (23.12)	.84
BMI 40-49.99 kg/m^2 at surgery, n (%)	8818 (54.96)	1243 (54.78)	7575 (55.00)	.84
BMI >50 kg/m^2 at surgery, n (%)	3016 (18.80)	441 (19.44)	2575 (18.69)	.40
Any lifetime cardiovascular disease event before surgery, n (%)	759 (4.73)	0 (0.00)	759 (5.51)	<.001
Gastroesophageal reflux disease in 2 years before surgery, n (%)	5799 (36.15)	800 (35.26)	4999 (36.29)	.34
Esophagitis in 2 years before surgery, n (%)	281 (1.75)	33 (1.45)	248 (1.80)	.24
Gastric ulcer in 2 years before surgery, n (%)	108 (0.07)	18 (0.08)	90 (0.07)	.45
Duodenal ulcer in 2 years before surgery, n (%)	1089 (6.79)	108 (4.76)	981 (7.12)	<.001
Peptic ulcer in 2 years before surgery, n (%)	264 (1.65)	39 (1.72)	225 (1.63)	.77
Gastritis duodenitis in 2 years before surgery, n (%)	1767 (11.01)	240 (10.58)	1527 (11.09)	.47
Dyspepsia in 2 years before surgery, n (%)	1837 (11.45)	262 (11.55)	1575 (11.43)	.88
Hiatal hernia in 2 years before surgery, n (%)	476 (2.97)	71 (3.13)	405 (2.94)	.62
Gastrointestinal bleed in 2 years before surgery, n (%)	9 (0.00)	2 (0.00)	7 (0.00)	.49
Aspirin use in 1 year before surgery, n (%)	2999 (18.69)	412 (18.16)	2587 (18.78)	.48
Aspirin use in 3 months before surgery, n (%)	1900 (11.84)	245 (10.80)	1655 (12.02)	.10
NSAID[a] use in 1 year before surgery, n (%)	6829 (42.57)	986 (43.46)	5843 (42.42)	.36
NSAID use in three months before surgery, n (%)	2342 (14.60)	346 (15.25)	1996 (14.49)	.34
Cirrhosis in 2 years before surgery, n (%)	84 (0.01)	12 (0.01)	72 (0.01)	.97
Sleep apnea in 2 years before surgery, n (%)	2330 (14.52)	346 (15.25)	1984 (14.40)	.29
Type 2 diabetes mellitus in 2 years before surgery, n (%)	5884 (36.68)	849 (37.42)	5035 (36.55)	.43

Variables	Accumulated 5 years of follow-up (N=16,043)	Missing 5-year data (N=2269)	Complete 5-year data (N=13,774)	*P* value
Hypertension in 2 years before surgery, n (%)	8768 (54.65)	1270 (55.97)	7498 (54.44)	.17
Chronic kidney disease in 2 years before surgery, n (%)	2170 (13.53)	319 (14.06)	1851 (13.44)	.42
Dyslipidemia in 2 years before surgery, n (%)	11,348 (70.73)	1844 (81.27)	9504 (69.00)	<.001
Any mental illness in 2 years before surgery, n (%)	8680 (54.10)	1304 (57.47)	7376 (5.355)	<.001
Attendance rate in 1 year before surgery (range 0%-100%), mean (SD)	76.70 (12.53)	76.50 (12.31)	76.70 (12.56)	.19
Any inpatient visit 1 year before surgery, n (%)	1081 (6.74)	159 (7.01)	922 (6.69)	.58
Any emergency department visit in 1 year before surgery, n (%)	3414 (21.28)	433 (19.08)	2981 (21.64)	.006

[a]NSAID: nonsteroidal anti-inflammatory drug.

Understanding the Decisions Between Bariatric Operations

Table 1 presents pairwise comparisons between VSG and RYGB patients in the ENGAGE CVD cohort to highlight the importance of using state-of-the-art statistical and econometric comparative effectiveness methods to adjust for differences in patient populations between those who receive VSG and those who have RYGB [23-26]. VSG patients, when compared with RYGB patients in the ENGAGE CVD cohort, were younger (*P*<.001), were more likely to be of a racial and ethnic minority group (*P*<.001), lost more weight before surgery (*P*<.001), and had a lower BMI (*P*<.001); and were less likely to have a BMI >50 kg/m^2 at the time of surgery (*P*<.001), had fewer lifetime CVD events (*P*=.002), and were less likely to be using aspirin before surgery (*P*<.001).

In addition, VSG patients when compared with RYGB patients in the ENGAGE CVD cohort had lower rates of GERD (*P*<.001), hiatal hernia (*P*=.003), sleep apnea (*P*<.001), T2DM (*P*<.001), hypertension (*P*<.001), chronic kidney disease (*P*<.001), dyslipidemia (*P*<.001), and mental illness (*P*<.001) at the time of surgery. Compared with RYGB patients, VSG patients had higher attendance rates for scheduled outpatient visits (*P*<.001) and lower rates of inpatient (*P*<.001) service use in the year before surgery.

As part of the process of understanding the decisions between bariatric operations, we conducted a series of meetings over 2 years with bariatric surgeons, patients, and providers about decisions they made between VSG and RYGB. We assembled a set of factors that our stakeholders felt were key determinants of why patients would undergo VSG or RYGB in Table 3. These factors were used to (1) construct propensity models with covariate adjustment and (2) test and select instrumental variables, which use natural variation to mimic random assignment to procedure, for comparative effectiveness analyses. Some of these variables, although important determinants of treatment assignment, were not included in our study because they were not available in the electronic health record. We included these variables in Table 3 because they illustrate the need to use statistical methods that can account for unmeasured confounders in the choice between bariatric operations. Most surgeons and providers indicated that patient preferences for one operation over another would be honored unless the operation they chose was a substantial safety risk for the patient.

Table 3. Factors considered as determinants in bariatric surgery decisions by a group of health care system stakeholders including patients, providers, and bariatric surgeons in the ENGAGE CVD (Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease) cohort study.

Factor	Preferred operation	Rationale	Available in electronic medical record
Year of surgery	Depends on year	Secular trends in surgery were apparent with RYGB[a] preferred in years before 2011 and VSG[b] preferred after 2011.	Yes
Preparation course instructor	Depends on instructor	Preparation course instructors have operation preferences and can communicate these to the patients and influence their choices.	Yes
Bariatric surgeon	Depends on surgeon	Surgeons have operation preferences as evidenced by frequency of type of operation over time.	Yes
Media consumption	Depends on source	Patients may be influenced to choose an operation based on electronic and other media consumption.	No
Patient race/ethnicity	VSG	More non-Hispanic black patients are having VSG compared with RYGB possibly because it is *less surgery*, and they will not lose *too much weight.*	Yes
History of cirrhosis and abdominal surgeries	VSG	Some bariatric surgeons believed that RYGB was inappropriate for patients with a history of cirrhosis and/or abdominal surgeries.	Yes
NSAID[c] and aspirin use	VSG	Some bariatric surgeons believed that patients requiring anti-inflammatories (NSAIDs, aspirin, and steroids) were high risk for surgery regardless of operation type; however, the highest risk was for RYGB.	Yes
BMI >50 kg/m^2	VSG	Some bariatric surgeons believed that much heavier patients had higher complication rates and that patients could be offered VSG to induce weight loss for a possible later, safer RYGB operation.	Yes
Medication-treated mental health	VSG	Some bariatric surgeons believed that patients requiring medication for mental health conditions may not do well after RYGB because of changes in absorption/metabolism after surgery.	Yes
Poor portion control	VSG	Some bariatric surgeons believed that if patients were severely obese mostly because of portion control, VSG would be the most conservative and successful option.	No
Complications	VSG	Most bariatric surgeons felt that VSG resulted in fewer complications than RYGB and should be the preferred operation to start, unless clearly contraindicated by GERD[d] or gastrointestinal conditions.	Yes
Sweet eating/craving	RYGB	Some bariatric surgeons believed that the adverse consequence of *dumping syndrome* with RYGB following sweet-eating binges was a good deterrent for these patients helping them be more successful.	No
Type 2 diabetes mellitus, hiatal hernia, and GERD	RYGB	Some bariatric surgeons believed that RYGB was better for diabetes remission, and hiatal hernia and GERD would complicate VSG.	Yes

[a]RYGB: Roux-en-Y gastric bypass.

[b]VSG: vertical sleeve gastrectomy.

[c]NSIAD: nonsteroidal anti-inflammatory drug.

[d]GERD: gastroesophageal reflux disease.

Discussion

Principal Findings

The ENGAGE CVD cohort was one of the largest sample sizes of real-world bariatric operations, especially VSG, which is now the most common operation performed in the United States [8]. Randomized controlled trials (RCTs) do not have the sample size necessary to properly explore HTE, which can guide subgroups of patients in their decision whether to choose weight loss surgery as a treatment option and then which operation to have [12,15]. In addition, the ENGAGE CVD cohort had an excellent long-term follow-up. Nearly 85.86% (13,774/16,043) of patients were members of the health care system 5 years after surgery (see Table 2). Finally, unlike almost any study in the bariatric surgical literature, the ENGAGE CVD cohort was 64% non-white, providing a rare opportunity to study the effects of bariatric surgery for different racial and ethnic minorities, some of whom have the highest rates of severe obesity in the United States [1]. The ENGAGE CVD cohort has a bariatric surgery patient profile similar to that of the United States in the next 5 to 10 years, as nationwide bariatric practice shifts strongly toward VSG and the United States becomes more racially and ethnically diverse.

Strengths and Weaknesses

The main weaknesses of the ENGAGE CVD cohort study were that all patients were insured, and although surgery was performed by 23 different surgeons across many settings, including surgeons outside of the health care system, the patients

in the ENGAGE CVD cohort were cared for primarily within a single integrated health care system. This health care system may not be representative of the care, both preoperatively and postoperatively, that other patients might receive in different kinds of health care settings. In addition, the data were assembled retrospectively from electronic health records that were designed for clinical care and not research. Thus, data were not systematically collected by research personnel at regular intervals. Outcomes were not assessed in a standardized way by research personnel and had to be defined using methods that combined the clinical information available in the electronic health record with clinical stakeholder input about treatment guidelines and practices. There were no mechanisms for obtaining measures from patients who missed appointments and/or disenrolled from the health care system. Despite these limitations, we have shown that data from electronic medical records, such as heights and weights, are valid and suitable for research [21].

In addition, patients were not randomly chosen for surgery from an eligible pool of participants and they were not randomly assigned to operations. This threatens both internal validity (differences between operations could have been because of the assignment process) and external validity (those receiving bariatric operations were not representative of all the patients who were potentially eligible to have these operations). RCTs would be the best statistical design to evaluate the causal *efficacy* of bariatric surgery for cardiovascular risk reduction (highest internal validity) [27,28]. However, RCTs have poor external validity and cannot answer questions about what will work in an uncontrolled real-world setting or in a population more heterogeneous than the restrictive trial sample that is typically studied [29]. Retrospective observational comparative effectiveness cohort studies such as ENGAGE CVD are better designs for testing how well *existing efficacious* treatments work for a heterogeneous patient population in an uncontrolled real-world setting.

Conclusions

The goal of the ENGAGE CVD study was to provide population-based, comprehensive, rigorous evidence for both clinical and policy decision making, informing the choice between RYGB and VSG for overall CVD risk reduction and risk factor remission, as well as safety in a diverse group of patients (racial and ethnic minority). Our findings will be used to provide recommendations to providers and patients about the decision between operations and help prioritize future health policy decisions and research investments in this area.

Acknowledgments

The ENGAGE CVD cohort study was funded by the National Heart, Lung, and Blood Institute award #1R01HL130462-01. We would like to thank our patients, providers and other health care system staff who care for them.

Conflicts of Interest

None declared.

References

1. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. J Am Med Assoc 2012 Feb 1;307(5):491-497. [doi: 10.1001/jama.2012.39] [Medline: 22253363]
2. Loveman E, Frampton GK, Shepherd J, Picot J, Cooper K, Bryant J, et al. The clinical effectiveness and cost-effectiveness of long-term weight management schemes for adults: a systematic review. Health Technol Assess 2011 Jan;15(2):1-182 [FREE Full text] [doi: 10.3310/hta15020] [Medline: 21247515]
3. Ribaric G, Buchwald JN, McGlennon TW. Diabetes and weight in comparative studies of bariatric surgery vs conventional medical therapy: a systematic review and meta-analysis. Obes Surg 2014 Mar;24(3):437-455 [FREE Full text] [doi: 10.1007/s11695-013-1160-3] [Medline: 24374842]
4. Schauer PR, Bhatt DL, Kirwan JP, Wolski K, Aminian A, Brethauer SA, STAMPEDE Investigators. Bariatric surgery versus intensive medical therapy for diabetes - 5-year outcomes. N Engl J Med 2017 Feb 16;376(7):641-651 [FREE Full text] [doi: 10.1056/NEJMoa1600869] [Medline: 28199805]
5. Maciejewski ML, Arterburn DE, van Scoyoc L, Smith VA, Yancy WS, Weidenbacher HJ, et al. Bariatric surgery and long-term durability of weight loss. JAMA Surg 2016 Nov 1;151(11):1046-1055 [FREE Full text] [doi: 10.1001/jamasurg.2016.2317] [Medline: 27579793]
6. Pollack A. New York Times.: New York Times; 2013 Jun 18. AMA Recognizes Obesity as a Disease New York URL: http://www.nytimes.com/2013/06/19/business/ama-recognizes-obesity-as-a-disease.html?hp&_r=0 [accessed 2020-02-07] [WebCite Cache ID 78tADaYMI]
7. Faria G. A brief history of bariatric surgery. Porto Biomed J 2017;2(3):90-92. [doi: 10.1016/j.pbj.2017.01.008]
8. Ponce J, Nguyen NT, Hutter M, Sudan R, Morton JM. American Society for Metabolic and Bariatric Surgery estimation of bariatric surgery procedures in the United States, 2011-2014. Surg Obes Relat Dis 2015;11(6):1199-1200. [doi: 10.1016/j.soard.2015.08.496] [Medline: 26476493]
9. Coleman KJ, Huang Y, Hendee F, Watson HL, Casillas RA, Brookey J. Three-year weight outcomes from a bariatric surgery registry in a large integrated healthcare system. Surg Obes Relat Dis 2014;10(3):396-403. [doi: 10.1016/j.soard.2014.02.044] [Medline: 24951065]

10. Fischer L, Hildebrandt C, Bruckner T, Kenngott H, Linke GR, Gehrig T, et al. Excessive weight loss after sleeve gastrectomy: a systematic review. Obes Surg 2012 May;22(5):721-731. [doi: 10.1007/s11695-012-0616-1] [Medline: 22411568]
11. Switzer NJ, Prasad S, Debru E, Church N, Mitchell P, Gill RS. Sleeve gastrectomy and type 2 diabetes mellitus: a systematic review of long-term outcomes. Obes Surg 2016 Jul;26(7):1616-1621. [doi: 10.1007/s11695-016-2188-y] [Medline: 27103028]
12. Trastulli S, Desiderio J, Guarino S, Cirocchi R, Scalercio V, Noya G, et al. Laparoscopic sleeve gastrectomy compared with other bariatric surgical procedures: a systematic review of randomized trials. Surg Obes Relat Dis 2013;9(5):816-829. [doi: 10.1016/j.soard.2013.05.007] [Medline: 23993246]
13. Yang X, Yang G, Wang W, Chen G, Yang H. A meta-analysis: to compare the clinical results between gastric bypass and sleeve gastrectomy for the obese patients. Obes Surg 2013 Jul;23(7):1001-1010. [doi: 10.1007/s11695-013-0938-7] [Medline: 23595210]
14. Chang SH, Stoll CR, Song J, Varela JE, Eagon CJ, Colditz GA. The effectiveness and risks of bariatric surgery: an updated systematic review and meta-analysis, 2003-2012. JAMA Surg 2014 Mar;149(3):275-287 [FREE Full text] [doi: 10.1001/jamasurg.2013.3654] [Medline: 24352617]
15. Li JF, Lai DD, Lin ZH, Jiang TY, Zhang AM, Dai JF. Comparison of the long-term results of Roux-en-Y gastric bypass and sleeve gastrectomy for morbid obesity: a systematic review and meta-analysis of randomized and nonrandomized trials. Surg Laparosc Endosc Percutan Tech 2014 Feb;24(1):1-11. [doi: 10.1097/SLE.0000000000000041] [Medline: 24487151]
16. Coleman KJ, Huang Y, Koebnick C, Reynolds K, Xiang AH, Black MH, et al. Metabolic syndrome is less likely to resolve in Hispanics and non-Hispanic blacks after bariatric surgery. Ann Surg 2014 Feb;259(2):279-285. [doi: 10.1097/SLA.0000000000000258] [Medline: 24100336]
17. DeMaria EJ, Pate V, Warthen M, Winegar DA. Baseline data from American Society for Metabolic and Bariatric Surgery - designated bariatric surgery centers of excellence using the bariatric outcomes longitudinal database. Surg Obes Relat Dis 2010;6(4):347-355. [doi: 10.1016/j.soard.2009.11.015] [Medline: 20176512]
18. Hutter MM, Schirmer BD, Jones DB, Ko CY, Cohen ME, Merkow RP, et al. First report from the American College of Surgeons Bariatric Surgery Center Network: laparoscopic sleeve gastrectomy has morbidity and effectiveness positioned between the band and the bypass. Ann Surg 2011 Sep;254(3):410-20; discussion 420 [FREE Full text] [doi: 10.1097/SLA.0b013e31822c9dac] [Medline: 21865942]
19. National Institutes for Diabetes and Digestive and Kidney Disorders (NIDDK). Potential Candidates for Bariatric Surgery URL: https://www.niddk.nih.gov/health-information/weight-management/bariatric-surgery/potential-candidates [accessed 2020-02-07]
20. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 2019 Jun 18;139(25):e1082-e1143. [doi: 10.1161/CIR.0000000000000625] [Medline: 30586774]
21. Arterburn DE, Ichikawa L, Ludman EJ, Operskalski B, Linde JA, Anderson E, et al. Validity of clinical body weight measures as substitutes for missing data in a randomized trial. Obes Res Clin Pract 2008 Dec;2(4):277-281 [FREE Full text] [doi: 10.1016/j.orcp.2008.09.002] [Medline: 19956347]
22. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014 Jun 24;129(25 Suppl 2):S49-S73. [doi: 10.1161/01.cir.0000437741.48606.98] [Medline: 24222018]
23. Basu A. Estimating person-centered treatment (PeT) effects using instrumental variables: An application to evaluating prostate cancer treatments. J Appl Econ (Chichester Engl) 2014;29(4):671-691 [FREE Full text] [doi: 10.1002/jae.2343] [Medline: 25620844]
24. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. Br Med J 2018 Dec 10;363:k4245 [FREE Full text] [doi: 10.1136/bmj.k4245] [Medline: 30530757]
25. Grieve R, O'Neill S, Basu A, Keele L, Rowan KM, Harris S. Analysis of benefit of intensive care unit transfer for deteriorating ward patients: a patient-centered approach to clinical evaluation. JAMA Netw Open 2019 Feb 1;2(2):e187704 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.7704] [Medline: 30768190]
26. Basu A, Chan KC. Can we make smart choices between OLS and contaminated IV methods? Health Econ 2014 Apr;23(4):462-472 [FREE Full text] [doi: 10.1002/hec.2926] [Medline: 23765683]
27. Salminen P, Helmiö M, Ovaska J, Juuti A, Leivonen M, Peromaa-Haavisto P, et al. Effect of Laparoscopic Sleeve Gastrectomy vs Laparoscopic Roux-en-Y Gastric Bypass on Weight Loss at 5 Years Among Patients With Morbid Obesity: The SLEEVEPASS Randomized Clinical Trial. JAMA 2018 Jan 16;319(3):241-254 [FREE Full text] [doi: 10.1001/jama.2017.20313] [Medline: 29340676]
28. Peterli R, Wölnerhanssen BK, Peters T, Vetter D, Kröll D, Borbély Y, et al. Effect of Laparoscopic Sleeve Gastrectomy vs Laparoscopic Roux-en-Y Gastric Bypass on Weight Loss in Patients With Morbid Obesity: The SM-BOSS Randomized Clinical Trial. J Am Med Assoc 2018 Jan 16;319(3):255-265 [FREE Full text] [doi: 10.1001/jama.2017.20897] [Medline: 29340679]

29. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med 2007 Aug;33(2):155-161. [doi: 10.1016/j.amepre.2007.04.007] [Medline: 17673104]

Abbreviations

CVD: cardiovascular disease
ENGAGE CVD: Effectiveness of Gastric Bypass versus Gastric Sleeve for Cardiovascular Disease
GERD: gastroesophageal reflux disease
HTE: heterogeneity of treatment effects
LIV: local instrumental variable
RCT: randomized controlled trial
RYGB: Roux-en-Y gastric bypass
T2DM: type 2 diabetes mellitus
VSG: vertical sleeve gastrectomy

Edited by G Eysenbach; submitted 05.06.19; peer-reviewed by A Notarnicola, MA Motamedi; comments to author 21.08.19; revised version received 08.10.19; accepted 16.12.19; published 06.04.20.

Please cite as:
Coleman KJ, Fischer H, Arterburn DE, Barthold D, Barton LJ, Basu A, Courcoulas A, Crawford CL, Fedorka P, Kim B, Mun E, Murali S, Reynolds K, Suh K, Wei R, Yoon TK, Zane R
Effectiveness of Gastric Bypass Versus Gastric Sleeve for Cardiovascular Disease: Protocol and Baseline Results for a Comparative Effectiveness Study
JMIR Res Protoc 2020;9(4):e14936
URL: https://www.researchprotocols.org/2020/4/e14936
doi:10.2196/14936
PMID:32249757

Protocol

Unibody Endograft Using AFX 2 for Less Invasive and Faster Endovascular Aortic Repair: Protocol for a Multicenter Nonrandomized Study

Roberto Silingardi[1*], MD; Pasqualino Sirignano[2*], MD; Francesco Andreoli[1], MD; Wassim Mansour[2], MD, PhD; Mattia Migliari[1], MD; Francesco Speziale[2], MD; LIVE Study Collaborators[3*]

[1]Department of Vascular Surgery, Ospedale Civile Sant'Agostino-Estense, Azienda Ospedaliero-Universitaria di Modena, University of Modena and Reggio Emilia, Modena, Italy

[2]Vascular and Endovascular Surgery Unit, Department of Surgery Paride Stefanini, Sapienza University of Rome, Rome, Italy

[3]See Acknowledgments section for collaborators/group members

[*]these authors contributed equally

Corresponding Author:
Pasqualino Sirignano, MD
Vascular and Endovascular Surgery Unit
Department of Surgery Paride Stefanini
Sapienza University of Rome
Viale del Policlinico, 155
Rome, 00161
Italy
Phone: 39 064940532
Email: pasqualino.sirignano@uniroma1.it

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2020/6/e20698/

Abstract

Background: Since the introduction of endovascular aortic repair (EVAR) for treatment of abdominal aortic aneurysms (AAAs), progressive improvements in results have been achieved. However, conventional bifurcated stent grafts have been proven to have a nonnegligible risk of failure and secondary intervention, principally due to the lack of adequate proximal sealing. The unique AFX 2 Endovascular AAA System (Endologix, Irvine, CA) unibody device, which provides different sealing and fixation features compared with conventional devices, seems to overcome these limitations.

Objective: The aim of this study is to evaluate intraoperative, perioperative, and postoperative results in patients treated with the AFX 2 Endovascular AAA System endografts for elective AAA repair in a large cohort of consecutive patients.

Methods: All eligible EVAR patients will be included in this observational, multicenter, prospective, nonrandomized study. The number of patients to be enrolled is 500.

Results: The primary endpoint of the study is to evaluate the technical and clinical success of EVAR with unibody endografts in short- (90-day), mid- (1-year), and long-term (5-year) follow-up periods. The following secondary endpoints will also be addressed: operative time, intraoperative radiation exposure, contrast medium usage, AAA sac shrinkage at 12-month and 5-year follow-up, and any potential role of patients' baseline characteristics and device configuration on primary endpoint. The actual start date of the investigation was November 2019. The final patient is expected to be treated by the end of December 2020, and the estimated study completion date is December 2025.

Conclusions: This study will provide verified real-world data on AAAs treated by AFX 2 endografts and followed for a long-term interval.

International Registered Report Identifier (IRRID): PRR1-10.2196/16959

(JMIR Res Protoc 2020;9(4):e16959) doi:10.2196/16959

KEYWORDS

aortic aneurysm; abdominal aortic aneurysm; endovascular aortic repair; endovascular repair; AFX 2; long-term results

Introduction

In recent years, endovascular aortic repair (EVAR) has emerged as a safe and valid option for treatment of abdominal aortic aneurysms (AAAs) (Multimedia Appendix 1). The AFX 2 Endovascular AAA System (Endologix, Irvine, CA) consists of two components: an implantable stent graft and a disposable delivery catheter. The preloaded stent graft is transferred through the AFX 2 Introducer System and inserted endoluminally via the femoral or iliac artery over a guidewire and, upon deployment and withdrawal of the delivery system, expands to the indicated diameter. During deployment and expansion, the stent graft is intended to form proximal and distal seal zones surrounding the aneurysm location. The stent graft is composed of a cobalt-chromium alloy self-expanding stent cage with a thin-walled, low porosity expanded polytetrafluoroethylene graft cover that is attached proximally and distally to the stent cage with a polypropylene suture. The system consists of a unibody bifurcated stent graft, with proximal extension and limb extension accessory components available, as needed, to accommodate the patient's specific anatomy.

The AFX 2 Endovascular AAA System is essentially composed of two distinct elements: the bifurcated stent graft and the proximal and iliac limb extension stent grafts.

The bifurcated element is the primary component that is delivered into the patient's aorta. All bifurcated stent grafts consist of a unibody configuration (an aortic main body with two attached iliac legs). The main body and each iliac leg are constructed from a single wire. The main body is manufactured in sizes ranging from 22 mm to 28 mm. The iliac legs are 13 mm to 20 mm in diameter for all sizes of bifurcated stent grafts.

The proximal and iliac limb extension stent graft components are used to extend the lengths of implanted bifurcated components. The AFX 2 Endovascular AAA System proximal extensions (Vela) are available in suprarenal and infrarenal configurations and use a circumferential radiopaque marker for identification of the proximal graft fabric line.

According to current instructions for use, the following anatomical criteria are required.

- Adequate iliac or femoral access compatible with the required delivery systems (diameter 6.5 mm)
- Nonaneurysmal aortic neck between the renal arteries and the aneurysm with a length 15 mm, a diameter ≥18 mm and ≤32 mm, and a neck angle ≤60° to the body of the aneurysm
- Aortic length ≥1.0 cm longer than the body portion of the chosen bifurcated model
- Common iliac artery distal fixation site with a distal fixation length ≥15 mm, ability to preserve at least one hypogastric artery, a diameter ≥10 mm and ≤23 mm, and an iliac angle ≤90° to the aortic bifurcation
- Extension stent grafts must have the ability to overlap the bifurcated stent graft by at least 30 to 40 mm proximally and at least 15 to 20 mm distally

Some elegant papers have been published on this unique device [1-7]. In 2010, Carpenter et al [6] reported in a study of 157 patients treated by unibody device implantation in three different prospective multicenter trials (Powerlink trial, Powerlink XL trial, and Powerlink Suprarenal Extension trial). All enrolled patients were treated between 2000 and 2008 and followed through 5 years. Technical success was achieved in 99% of patients. Aneurysm exclusion was achieved in all patients over a mean procedure time of 132 (SD 58) minutes. No aneurysm related deaths, ruptures, conversions, or migrations have been observed to current follow-up, as these aneurysms have continued to remodel with more than 92% of patients free of sac growth. At each annual evaluation period, no stent fractures, material failures, or losses of patency were found by the core laboratory. During the follow-up period, 5 patients were treated for a Type Ia endoleak, 3 for a Type Ib, and 3 for a limb occlusion. All of these reinterventions were performed with a new endovascular procedure without needing surgical conversion [6].

Similar results were reported by Qu and Raithel [4] in their single center study on more than 600 patients. Among the 612 patients in the cohort, 99 cases (16%) completed between 1999 and 2004, had the endograft deployed from the renal artery downward. The remaining 513 (84%) had the bifurcated stent graft deployed onto the native bifurcation, and among those cases 146 (28%) were deemed as challenging anatomy with a short or angulated neck. Technical success was achieved in 98.5% of patients (603/612). Intraoperative conversion occurred in 9 patients: 8 delivery access failures and 1 rupture. Perioperatively, 3 deaths occurred, and 2 limb occlusions were encountered. The rates of late conversion in the renal fixation and anatomical fixation groups were 4.0% and 1.9%, respectively. Likewise, the cumulative rates of a type I proximal endoleak in the renal fixation and anatomical fixation groups were 5.0% and 1.2%, respectively. Remarkably, no stent fracture, graft disruption, or type III or type IV endoleak was observed in their experience. Freedom from aneurysm sac diameter increase was 96% [4].

Moreover, Silingardi et al [7] performed a comparative study to compare nephrotoxic contrast medium with radiation exposure during elective EVAR procedures using unibody and modular devices. The initial hypothesis was that the unique unibody device structure, associated with not needing a gate cannulation, could reduce total procedural and total fluoroscopy time, as well as reduce the volume of contrast medium needed. Their study on 60 unibody devices and 57 bifurcated devices confirmed the hypothesis. For unibody and bifurcated devices, the median surgical procedure duration was 75 min vs 105 min (P<.001), the median volume of iodine contrast injected was 85 ml vs 170 ml (P<.001), and the median fluoroscopy time was 350 sec vs 780 sec (P<.001), respectively [7].

The promising data from these studies should be confirmed by prospective data collecting in a large consecutive cohort of patients using the latest generation unibody device implantation.

Therefore, this study aims to evaluate intraoperative, perioperative, and postoperative results in patients treated with AFX 2 Endovascular AAA System endografts for elective AAA repair in a large cohort of consecutive patients.

Methods

Objective and Duration of Investigation

The aim of this study is to evaluate intraoperative, perioperative, and postoperative results in patients treated using the latest generation AFX 2 Endovascular AAA System endograft for elective AAA repair in a multicentric study.

A total of 46 south European high-volume centers across Italy and Spain were involved in the Less Invasive and Faster Endovascular Aortic Repair Study. In a 12-month period from October 2018 to October 2019, mean EVAR procedures per center were 49.56 (range 20-140), while mean AFX procedures per year per center were 14.54 (range 10-57).

All consecutive eligible patients submitted to EVAR by AFX 2 Endovascular AAA System implantation will be included in the analysis. Patients will be submitted to EVAR procedures on the basis of their own preferences, anatomical features, and the operator's experience.

In light of the participating centers' numbers and activity volumes, an estimated 500 patients submitted to EVAR with AFX 2 should be enrolled. The sample size is low enough to obtain statistically significative results, according to previous published studies [1-7]. The anticipated duration of this clinical investigation is approximately 6 years. It is estimated that the inclusion period will be 12 months. The follow-up period is set to be 5 years.

Prior to enrollment in the clinical investigation, patients will be evaluated by their physician for the inclusion criteria. Each patient's medical condition should be stable, with no underlying medical condition that would prevent them from performing the required testing or completing the study. Patients should be geographically stable, willing and able to cooperate in this clinical study, and remain available for a midterm follow-up. Patients who do not wish to participate in this study can obtain the best available EVAR therapy as indicated, that is refusal to participate in this study will in no way affect their care at the institution. Inclusion and exclusion criteria are detailed in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for the study.

Inclusion criteria

- Elective abdominal aortic aneurysm patients that should be treated by standard abdominal endovascular aneurysm repair, according to Endologix AFX unibody device's instructions for use
- Patient is willing to comply with specified follow-up evaluations at the specified times for the duration of the study
- Patient is >18 years of age
- Patient, or their legal representative, understands the nature of the procedure and provides written informed consent prior to enrollment in the study

Exclusion criteria

- Abdominal endovascular aneurysm repair performed in urgent or emergent setting
- Patients treated outside Endologix AFX unibody device's instructions for use
- Patients refusing treatment
- Patients for whom antiplatelet therapy, anticoagulants, or antihypertensive drugs are contraindicated
- Patients with a history of prior life-threatening contrast medium reaction
- Life expectancy is less than follow-up period

This study respects all the principles reported in the current version of the Helsinki declaration (2013). According to the current International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use and Good Clinical Practice guidelines, each investigator is responsible for the regularity of the study. The aim of these standards is to assure the safety and comfort of all people recruited in the study. The study protocol and the written informed consensus form will be submitted to the local ethics committees for review.

AAA morphology will be assessed by OsiriX MD (PIXMEO, Geneva, Switzerland) on a regular Mac OS computer in one preoperative, contrast-enhanced, computed tomography angiography (CTA) [8]. The CTA must be performed with a biphasic acquisition protocol (unenhanced and contrast-enhanced scanning with a bolus tracking system) and reconstructions of 1-mm slices. All measurements (diameter, length, and angle) will be evaluated on a workstation with dedicated reconstruction software for center lumen line analysis and multiplanar reconstruction.

A patient is considered enrolled in the study if there is full compliance with the study inclusion and exclusion criteria.

Clinical data will be collected at patient enrollment, the EVAR procedure, discharge, planned follow-ups (ie, 1-3 months and 12 months postprocedure, and yearly thereafter), unplanned or interim follow-ups, and patient death. CTAs are mandatory within 90 days and then at 1 and 5 years after the index procedure. The duplex ultrasound scan will be performed at the

same follow-up interval and also at 24, 36, and 48 months. A new CTA will be performed in case of unexpected events during follow-up.

Endpoints

The primary endpoint of the study is to evaluate the technical and clinical success of EVAR with unibody endografts in short- (90-day), mid- (1-year), and long-term (5-year) follow-up periods.

Technical success was defined as the correct graft deployment without any unintentional occlusion of the aortic visceral branches or both hypogastric arteries, with aneurysm exclusion confirmed by the intraoperative angiography, no signs of type I or III endoleak, or conversion to open surgery.

Clinical success included successful deployment of the endovascular device at the intended location without death as a result of aneurysm-related treatment, type I or III endoleak, graft infection or thrombosis, aneurysm expansion (>5 mm), aneurysm rupture, or conversion to open repair (OR), as well as the presence of graft dilatation of 20% or more by diameter, graft migration, or a failure of device integrity [9].

The clinical and technical success were defined as assisted in cases of unplanned endovascular procedures, or secondary if unplanned surgery is necessary.

The following secondary endpoints will be also addressed: operative time, intraoperative radiation exposure, contrast medium usage, AAA sac shrinkage at 12-month and 5-year follow-ups, and any potential role of patients' baseline characteristics and device configuration on primary endpoint.

Data Collection and Analysis

Patient data will be captured electronically using a computer-based platform accessible to all investigators. Descriptive data summaries will be used to present and summarize the collected data. For categorical variables such as gender, frequency distributions and cross tabulations will be given. For numeric variables such as patient age, minimum, maximum, mean, median, and standard deviation will be calculated. For all variables, a 95% confidence interval for the relevant parameters of the underlying distribution will be calculated. For all time-dependent events, life tables will be calculated using the Kaplan-Meier estimate method for a period starting on the date of the procedure up to and including all follow-up visits. Stratification to risk factors will be performed and the logrank test will be used to compare the different outcomes; associated $P<.05$ will be defined as significant.

All preoperative and follow-up CTAs were assessed and independently evaluated by two experienced vascular surgeons at core lab centers. Disagreements will be discussed and resolved by consensus.

Patient Confidentiality

All information and data concerning patients or their participation in this clinical investigation will be considered confidential. Only authorized personnel will have access to these confidential files. Authorized personnel of health authorities will have the right to inspect and copy all records pertinent to this clinical investigation. All data used in the analysis and reporting of this clinical investigation will be anonymized.

Results

The actual start date of the investigation was November 2019. It is anticipated that 500 patients will be recruited to the study. The final patient is expected to be treated by the end of December 2020 and the estimated study completion date is December 2025. After data analysis, results will be shared with each investigator.

Discussion

In the last years, EVAR has become the standard of care for AAA treatment, and nowadays it represents the recommended modality of treatment according to the European Society for Vascular and Endovascular Surgery and the Society for Vascular Surgery guidelines [10,11].

However, the major randomized controlled trials (RCTs) on EVAR vs OR have not reached definitive conclusions. In 2004, the results of the first two RCTs were published [12,13].

The EVAR-1 trial described a clear advantage of EVAR compared to OR at 30 days. Greenhalgh and collaborators [12] reported that 30-day mortality in the EVAR group was 1.7% (9/531) vs 4.7% (24/516) in the OR group ($P=.009$). Their results were interpreted as a license to continue scientific evaluation of EVAR, but not to change clinical practice [12].

More enthusiastic conclusions came from the analysis of the Dutch Randomised Endovascular Aneurysm Management (DREAM) trial, reporting an operative mortality rate of 4.6% in the OR group (8/174 patients) and 1.2% in the EVAR group (2/171 patients) in a series of patients treated between 2000 and 2003, resulting in a risk ratio of 3.9 (95% CI 0.9-32.9). The combined rate of operative mortality and severe complications was 9.8% in the OR group (17/174 patients) and 4.7% in the EVAR group (8/171 patients), resulting in a risk ratio of 2.1 (95% CI 0.9-5.4). The authors concluded that EVAR was preferable to OR [13].

However, long-term follow-up was demanded to determine whether advantages persisted, and 1 year later, both trials published their midterm results [14,15]. In the EVAR-1 trial, all-cause mortality was similar in the two groups (approximately 28%; $P=.46$), but a persistent reduction in AAA-related deaths was recorded in the EVAR group compared with the OR group (4% vs 7%; $P=.04$). On the other hand, the proportion of patients with postoperative complications was 41% in the EVAR group and 9% in the OR group ($P<.001$) [14]. In the DREAM trial, 2 years after randomization, the cumulative survival rates were 89.6% for OR and 89.7% for EVAR and the cumulative rates of AAA-related deaths were 5.7% for OR and 2.1% for EVAR. This advantage of EVAR over OR was accounted for by events occurring in the perioperative period, with no significant difference in subsequent AAA-related mortality. The rate of survival free of moderate or severe complications was also

similar in the two groups at 2 years (65.9% OR vs 65.6% EVAR) [15].

The Open Versus Endovascular Repair (OVER) trial, which included patients treated between 2002 and 2007, was published in 2010. On the basis of a mean follow-up of 1.8 years, the OVER trial results showed that perioperative mortality was lower for EVAR than for OR (0.5% vs 3.0%; *P*=.004), without any difference at 2 years (7.0% vs 9.8%, *P*=.13). Mortality after the perioperative period was similar in the two groups (6.1% vs 6.6%); however, 4 late deaths in the EVAR group were AAA-related compared with none in the OR group. No differences between the two groups in terms of major morbidity, procedure failure, secondary intervention, AAA-related hospitalization, or health-related quality of life were recorded. Interestingly, no increase in midterm mortality after EVAR resulted in the loss of early survival advantage as shown in previous trials was observed [16].

A year later, a French RCT (ACE Trial) reported quite different results with no differences between EVAR and OR. Although only low to intermediate risk patients were enrolled, OR and EVAR offered no difference in survival (96.5% vs 95.2% at 1 year, and 86.7% vs 86.3% at 3 years) or in major and minor complications (95.9% vs 93.2% at 1 year, and 85.1% vs 82.4% at 3 years) [17].

These results led to a change in point of view: EVAR was considered as feasible as OR without any advantages, even in the short-term. The same year, a new US study, with a 6-year follow-up on 45,652 Medicare beneficiaries undergoing EVAR or OR in the period between 2001 and 2004, was analyzed to clarify the late results of endovascular procedures. Throughout follow-up, overall reintervention or readmission rates were similar with the two repair methods but slightly more common after EVAR than OR (7.6 vs 7.0/100 person-years; *P*<.001). Overall 30-day mortality with any reintervention or readmission was 9.1%. EVAR patients had more AAA-related reinterventions than OR (3.7% vs 0.9%; *P*<.001; mortality 5.6%). Conversely, EVAR patients had fewer laparotomy-related reinterventions than OR patients (1.4% vs 3.0%; *P*<.001; mortality, 8.1%) and fewer readmissions without surgery (2.0% vs 2.7%; *P*<.001; mortality 10.9%). Overall, reinterventions and readmissions accounted for 9.6% of all EVAR deaths and 7.6% of all OR deaths in the follow-up period (*P*<.001). The authors concluded that reintervention and readmission were slightly higher after EVAR. Survival was negatively affected by reintervention or readmission after EVAR and OR [18].

In 2016, the long-term results of the EVAR-1 RCT were published. In a mean of 12.7 years for follow-up, Patel et al [19] reported a similar overall mortality in the EVAR and OR groups (9.3 deaths per 100 person-years vs 8.9 deaths per 100 person-years; hazard ratio (HR) 1.11, 95% CI 0.97-1.27). However, a time analysis showed that beyond 8 years after randomization OR had a significantly lower mortality (HR 1.25, 95% CI 1.00-1.56, for total mortality; and HR 5.82, 95% CI 1.64-20.65, *P*=.006 for aneurysm-related mortality) mainly due to secondary sac rupture. The authors concluded that the early benefits of EVAR in terms of mortality were lost in the long-term [19].

However, late results from the OVER trial were published. In Lederle's [20] study, long-term overall survival was similar between patients who underwent endovascular repair and those who underwent OR (EVAR 68% vs OR 70%; HR 0.96, 95% CI 0.82-1.13). A difference between groups was noted in the number of patients who underwent secondary therapeutic procedures (EVAR 26.7% vs OR 19.8; 95% CI 2.0-17.5). Their results were not consistent with the findings of worse performance of endovascular repair with respect to long-term survival that was seen in the two European trials [20].

Notably, all these trials reported only few data or none at all on patients treated by unibody stent graft implantation. As mentioned above, AFX 2 Endovascular AAA System endografts are completely different from a technical and philosophical point of view from all other modular devices, and it seems to also provide different results. In fact, different studies have already demonstrated the advantages of this endograft and its safety and efficacy in short and midterm follow-up periods [1-5]. The preservation of the aortic bifurcation focuses the disrupting forces caused by columnar blood on the carrefour rather than on the aortic neck minimizing migration risk [1,2]. The obviation of contralateral gate cannulation makes this graft faster to deploy and uniquely suitable in cases of narrow aortic bifurcations. Moreover, this graft is allowed for future contralateral access for lower extremity interventions in a patient population with different vascular diseases. Finally, Dietrich et al [5] affirmed that the big contact area provided by the fabric free to move with the blood pressure wave can promote sac shrinkage and contrast type II endoleak formations.

Given the lack in current literature of effective data on unibody endograft results, the aim of our prospective study is addressing intraoperative, perioperative, and postoperative results in patients treated by the latest generation AFX 2 Endovascular AAA System endografts for elective AAA repair in a multicentric study.

Acknowledgments

The LIVE Study Collaborators are as follows: Giancarlo Accarino; Dimitri Apostolou; Guido Bajardi; Stefano Bartoli; Filippo Benedetto; Franco Briolini; Stefano Camparini; Emidio Costantini; Giovanni Credi; Ruggiero Curci; Raffaello Dallatana; Gianmarco de Donato; Carlo Dionisi; Vittorio Dorrucci; Leonardo Ercolini; Gianfranco Fadda; Mauro Ferrari; Loris Flora; Andrea Gaggiano; Roberto Gattuso; Franco Grego; Sabrina Grimaldi; Giovanni Impedovo; Arnaldo Ippoliti; Antonio Jannello; Sergio Losa; Nicola Mangialardi; Isaac Martinez; Javier Martinez; Stefano Michelagnoli; Giancarlo Palasciano; Vincenzo Palazzo; Domenico Palombo; Raffaele Pulli; Sonia Ronchey; Giovanni Rossi; Antonino Scolaro; Gianantonio Simoni; Francesco Spinelli;

Francesco Talarico; Maurizio Taurino; Marco Trogolo; Nicola Tusini; Gianfranco Veraldi; Pier Francesco Veroux; Gennaro Vigliotti; and Pietro Volpe.

Conflicts of Interest

None declared.

Multimedia Appendix 1
EVAR Overview.
[DOCX File , 20 KB - resprot_v9i4e16959_app1.docx]

References

1. Coppi G, Silingardi R, Tasselli S, Gennai S, Saitta G, Veraldi GF. Endovascular treatment of abdominal aortic aneurysms with the Powerlink Endograft System: influence of placement on the bifurcation and use of a proximal extension on early and late outcomes. J Vasc Surg 2008 Oct;48(4):795-801 [FREE Full text] [doi: 10.1016/j.jvs.2008.05.011] [Medline: 18586447]
2. Wang GJ, Carpenter JP, Endologix Investigators. The Powerlink system for endovascular abdominal aortic aneurysm repair: six-year results. J Vasc Surg 2008 Sep;48(3):535-545 [FREE Full text] [doi: 10.1016/j.jvs.2008.04.031] [Medline: 18635335]
3. Albertini J, Lahlou Z, Magnan P, Branchereau A, French Powerlink Multicenter Trial Investigators. Endovascular repair of abdominal aortic aneurysms with a unibody stent-graft: 3-year results of the French Powerlink Multicenter Trial. J Endovasc Ther 2005 Dec;12(6):629-637. [doi: 10.1583/05-1629R.1] [Medline: 16363890]
4. Qu L, Raithel D. From clinical trials to clinical practice: 612 cases treated with the Powerlink stent-graft for endovascular repair of AAA. J Cardiovasc Surg (Torino) 2009 Apr;50(2):131-137. [Medline: 19329908]
5. Diethrich EB. Novel sealing concept in the Endologix AFX unibody stent-graft. J Cardiovasc Surg (Torino) 2014 Feb;55(1):93-102. [Medline: 24356051]
6. Carpenter JP, Garcia MJ, Harlin SA, Jordan WD, Jung MT, Krajcer Z, et al. Contemporary results of endovascular repair of abdominal aortic aneurysms: effect of anatomical fixation on outcomes. J Endovasc Ther 2010 Apr;17(2):153-162. [doi: 10.1583/09-2977.1] [Medline: 20426630]
7. Silingardi R, Azzoni I, Giuliani E, Saitta G, Gennai S, Coppi G. Unibody endografts for abdominal aortic aneurysm repair reduce radiation and nephrotoxic exposure compared with modular endografts. Ann Vasc Surg 2015;29(4):751-757. [doi: 10.1016/j.avsg.2014.11.011] [Medline: 25637576]
8. Setacci F, Sirignano P, Cappelli A, Setacci C. The wonders of a newly available post-analysis CT software in the hands of vascular surgeons. Eur J Vasc Endovasc Surg 2012 Apr;43(4):404-406 [FREE Full text] [doi: 10.1016/j.ejvs.2011.11.027] [Medline: 22226699]
9. Chaikof EL, Blankensteijn JD, Harris PL, White GH, Zarins CK, Bernhard VM, Ad Hoc Committee for Standardized Reporting Practices in Vascular Surgery of The Society for Vascular Surgery/American Association for Vascular Surgery. Reporting standards for endovascular aortic aneurysm repair. J Vasc Surg 2002 May;35(5):1048-1060 [FREE Full text] [doi: 10.1067/mva.2002.123763] [Medline: 12021727]
10. Wanhainen A, Verzini F, Van Herzeele I, Allaire E, Bown M, Cohnert T, Esvs Guidelines Committee, et al. Editor's choice - European Society for Vascular Surgery (ESVS) 2019 clinical practice guidelines on the management of abdominal aorto-iliac artery aneurysms. Eur J Vasc Endovasc Surg 2019 Jan;57(1):8-93 [FREE Full text] [doi: 10.1016/j.ejvs.2018.09.020] [Medline: 30528142]
11. Chaikof EL, Dalman RL, Eskandari MK, Jackson BM, Lee WA, Mansour MA, et al. The Society for Vascular Surgery practice guidelines on the care of patients with an abdominal aortic aneurysm. J Vasc Surg 2018 Jan;67(1):2-77.e2 [FREE Full text] [doi: 10.1016/j.jvs.2017.10.044] [Medline: 29268916]
12. Greenhalgh RM, Brown LC, Kwong GPS, Powell JT, Thompson SG, EVAR trial participants. Comparison of endovascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: randomised controlled trial. Lancet 2004;364(9437):843-848. [doi: 10.1016/S0140-6736(04)16979-1] [Medline: 15351191]
13. Prinssen M, Verhoeven ELG, Buth J, Cuypers PWM, van Sambeek MRHM, Balm R, Dutch Randomized Endovascular Aneurysm Management (DREAM) Trial Group. A randomized trial comparing conventional and endovascular repair of abdominal aortic aneurysms. N Engl J Med 2004 Oct 14;351(16):1607-1618. [doi: 10.1056/NEJMoa042002] [Medline: 15483279]
14. EVAR trial participants. Endovascular aneurysm repair versus open repair in patients with abdominal aortic aneurysm (EVAR trial 1): randomised controlled trial. Lancet 2005;365(9478):2179-2186. [doi: 10.1016/S0140-6736(05)66627-5] [Medline: 15978925]
15. Blankensteijn JD, de Jong SECA, Prinssen M, van der Ham AC, Buth J, van Sterkenburg SMM, Dutch Randomized Endovascular Aneurysm Management (DREAM) Trial Group. Two-year outcomes after conventional or endovascular

repair of abdominal aortic aneurysms. N Engl J Med 2005 Jun 09;352(23):2398-2405. [doi: 10.1056/NEJMoa051255] [Medline: 15944424]
16. Lederle FA, Freischlag JA, Kyriakides TC, Padberg FT, Matsumura JS, Kohler TR, Open Versus Endovascular Repair (OVER) Veterans Affairs Cooperative Study Group. Outcomes following endovascular vs open repair of abdominal aortic aneurysm: a randomized trial. JAMA 2009 Oct 14;302(14):1535-1542. [doi: 10.1001/jama.2009.1426] [Medline: 19826022]
17. Becquemin J, Pillet J, Lescalie F, Sapoval M, Goueffic Y, Lermusiaux P, et al. A randomized controlled trial of endovascular aneurysm repair versus open surgery for abdominal aortic aneurysms in low- to moderate-risk patients. J Vasc Surg 2011 May;53(5):1167-1173.e1 [FREE Full text] [doi: 10.1016/j.jvs.2010.10.124] [Medline: 21276681]
18. Giles KA, Landon BE, Cotterill P, O'Malley AJ, Pomposelli FB, Schermerhorn ML. Thirty-day mortality and late survival with reinterventions and readmissions after open and endovascular aortic aneurysm repair in Medicare beneficiaries. J Vasc Surg 2011 Jan;53(1):6-12,13.e1 [FREE Full text] [doi: 10.1016/j.jvs.2010.08.051] [Medline: 21030195]
19. Patel R, Sweeting MJ, Powell JT, Greenhalgh RM, EVAR Trial Investigators. Endovascular versus open repair of abdominal aortic aneurysm in 15-years' follow-up of the UK endovascular aneurysm repair trial 1 (EVAR trial 1): a randomised controlled trial. Lancet 2016 Nov 12;388(10058):2366-2374 [FREE Full text] [doi: 10.1016/S0140-6736(16)31135-7] [Medline: 27743617]
20. Lederle FA, Kyriakides TC, Stroupe KT, Freischlag JA, Padberg FT, Matsumura JS, OVER Veterans Affairs Cooperative Study Group. Open versus endovascular repair of abdominal aortic aneurysm. N Engl J Med 2019 May 30;380(22):2126-2135. [doi: 10.1056/NEJMoa1715955] [Medline: 31141634]

Abbreviations

AAA: abdominal aortic aneurysm
CTA: computed tomography angiography
DREAM: Dutch Randomized Endovascular Aneurysm Management
EVAR: endovascular aortic repair
HR: hazard ratio
OR: open repair
OVER: Open Versus Endovascular Repair
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 07.11.19; peer-reviewed by V Brizzi, N Montelione; comments to author 10.01.20; revised version received 20.01.20; accepted 26.01.20; published 06.04.20.

Please cite as:
Silingardi R, Sirignano P, Andreoli F, Mansour W, Migliari M, Speziale F, LIVE Study Collaborators
Unibody Endograft Using AFX 2 for Less Invasive and Faster Endovascular Aortic Repair: Protocol for a Multicenter Nonrandomized Study
JMIR Res Protoc 2020;9(4):e16959
URL: http://www.researchprotocols.org/2020/4/e16959/
doi:10.2196/16959
PMID:32250278

Protocol

Investigating Health Impacts of Natural Resource Extraction Projects in Burkina Faso, Ghana, Mozambique, and Tanzania: Protocol for a Mixed Methods Study

Andrea Farnham[1,2], MPH, PhD; Hermínio Cossa[1,3], BA, MSc; Dominik Dietler[1,2], BA, MSc; Rebecca Engebretsen[4], BA, MSc, PhD; Andrea Leuenberger[1,2], BA, MSc; Isaac Lyatuu[1,2,5], BA, MSc; Belinda Nimako[1,2,6], BA, MD; Hyacinthe R Zabre[1,2,7], BA, MPH; Fritz Brugger[4], BA, PhD; Mirko S Winkler[1,2], MSc, PhD, DTM&H

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

[3]Manhiça Health Research Centre, Maputo, Mozambique

[4]Swiss Federal Institute of Technology, Zurich, Switzerland

[5]Ifakara Health Institute, Dar es Salaam, United Republic of Tanzania

[6]University of Health and Allied Sciences, Ho, Ghana

[7]Research Institute of Health Sciences, Ouagadougou, Burkina Faso

Corresponding Author:
Andrea Farnham, MPH, PhD
Swiss Tropical and Public Health Institute
PO Box
Basel, 4002
Switzerland
Phone: 41 612848684
Email: andrea.farnham@swisstph.ch

Abstract

Background: Natural resource extraction projects offer both opportunities and risks for sustainable development and health in host communities. Often, however, the health of the community suffers. Health impact assessment (HIA) can mitigate the risks and promote the benefits of development but is not routinely done in the developing regions that could benefit the most.

Objective: Our study aims to investigate health and health determinants in regions affected by extractive industries in Burkina Faso, Ghana, Mozambique, and Tanzania. The evidence generated in our study will inform a policy dialogue on how HIA can be promoted as a regulatory approach as part of the larger research initiative called the HIA4SD (Health impact assessment for sustainable development) project.

Methods: The study is a concurrent triangulation, mixed methods, multi-stage, multi-focus project that specifically addresses the topics of governance and policy, social determinants of health, health economics, health systems, maternal and child health, morbidity and mortality, and environmental determinants, as well as the associated health outcomes in natural resource extraction project settings across four countries. To investigate each of these health topics, the project will (1) use existing population-level databases to quantify incidence of disease and other health outcomes and determinants over time using time series analysis; (2) conduct two quantitative surveys on mortality and cost of disease in producer regions; and (3) collect primary qualitative data using focus groups and key informant interviews describing community perceptions of the impacts of extraction projects on health and partnership arrangements between the projects and local and national governance. Differences in health outcomes and health determinants between districts with and without an extraction project will be analyzed using matched geographical analyses in quasi-Poisson regression models and binomial regression models. Costs to the health system and to the households from diseases found to be associated with projects in each country will be estimated retrospectively.

Results: Fieldwork for the study began in February 2019 and concluded in February 2020. At the time of submission, qualitative data collection had been completed in all four study countries. In Burkina Faso, 36 focus group discussions and 74 key informant interviews were conducted in three sites. In Ghana, 34 focus group discussions and 64 key informant interviews were conducted in three sites. In Mozambique, 75 focus group discussions and 103 key informant interviews were conducted in four sites. In Tanzania, 36 focus group discussions and 84 key informant interviews were conducted in three sites. Quantitative data extraction

and collection is ongoing in all four study countries. Ethical approval for the study was received in all four study countries prior to beginning the fieldwork. Data analyses are underway and results are expected to be published in 2020 and 2021.

Conclusions: Disentangling the complex interactions of resource extraction projects with their host communities requires an integrative approach drawing on many methodologies under the HIA umbrella. By using complementary data sources to address the question of population health in project areas from several angles, bias and missing data will be reduced, generating high-quality evidence to aid countries in moving toward sustainable development.

International Registered Report Identifier (IRRID): DERR1-10.2196/17138

(JMIR Res Protoc 2020;9(4):e17138) doi:10.2196/17138

KEYWORDS

health impact assessment; environmental health; mixed methods; extractive industry; DHIS2; time series; GIS; cost-benefit; DHS

Introduction

Background

Natural resource extraction projects (eg, minerals, metals, oil, and gas) are major drivers of the economy in many developing countries, offering opportunities for sustainable economic and social growth for the local population, with accompanying implications for human health [1,2]. The nature of large natural resource extraction projects often causes upheaval in many sectors that affect both health and health determinants, including public health, society, and ecosystems [3-6]. Projects move in and offer new opportunities for jobs, improved infrastructure, a strengthened health system, and economic mobility [7]. However, historically the development of these projects is often instead accompanied by negative social and health outcomes in the surrounding community, often termed “the resource curse” [8-11]. These negative outcomes have included environmental contamination, strain on local water and sanitation resources from in-migration, reduced health equity, accidents and injuries caused by increased traffic, and increases in sexually transmitted, vector-borne, and chronic diseases [4-6,8,10,12-16]. The complex nature of resource extraction projects means that, in this upheaval, an integrated approach to identify potential opportunities and risks for the health of the population is needed.

Health impact assessment (HIA) is one approach that has the potential to harness the economic and social potential of extractive industries for the benefit of the local population. HIA is a method of impact assessment that uses a combination of quantitative and qualitative methods to prospectively identify the potential positive and negative health impacts of projects, policies, or programs [17]. In addition to investigating direct impacts on health, such as changes in disease and accident incidence, HIA gathers evidence on indirect impacts on the social, environmental, and institutional determinants of health, such as access to water and sanitation, food availability, and health system capacity, generating a strong evidence base from which decisions about public health policy and management can be made [18]. Despite the fact that environmental impact assessment has become accepted practice for mitigating environmental risks during project implementation [19,20], the use of HIA lags far behind [21]. Large gaps thus remain in our understanding of the complex picture of health and health equity in regions impacted by resource extraction projects. This need is particularly pressing in regions such as sub-Saharan Africa that already have low scores on the health-related targets of the Sustainable Development Goals (SDGs) [22]. This represents a wasted opportunity to turn the massive investments of the extractive industry into a net positive for the local population by minimizing negative health impacts of project implementation and maximizing the effects of the positive inflow of new resources on local and national development.

Study Objectives

Our project aims to use the HIA toolbox across natural resource extraction project settings in Burkina Faso, Ghana, Mozambique, and Tanzania, countries acutely affected by the development of extractive industries. By doing so, the study will generate new evidence about health in regions affected by resource extraction and translate these findings into actionable policy recommendations. In addition, by partnering with local institutions and training local PhD students in HIA methodology, the project will build long-term capacity for carrying out HIA in the low- and middle-income countries where it is most needed.

Specifically, the study aims to (1) evaluate the effects of natural resource extraction projects on health-related targets and associated indicators of the SDGs that can be observed at the national and local level; (2) assess how projects interact with, and have effects on, local health systems; (3) determine the costs and benefits of projects for local and national health systems; and (4) characterize the influence projects exert on health equity in affected populations. The evidence generated in the study will inform a policy dialogue on whether and how HIA can be promoted as a regulatory approach as part of the larger research initiative called the HIA4SD (Health impact assessment for sustainable development) project [23].

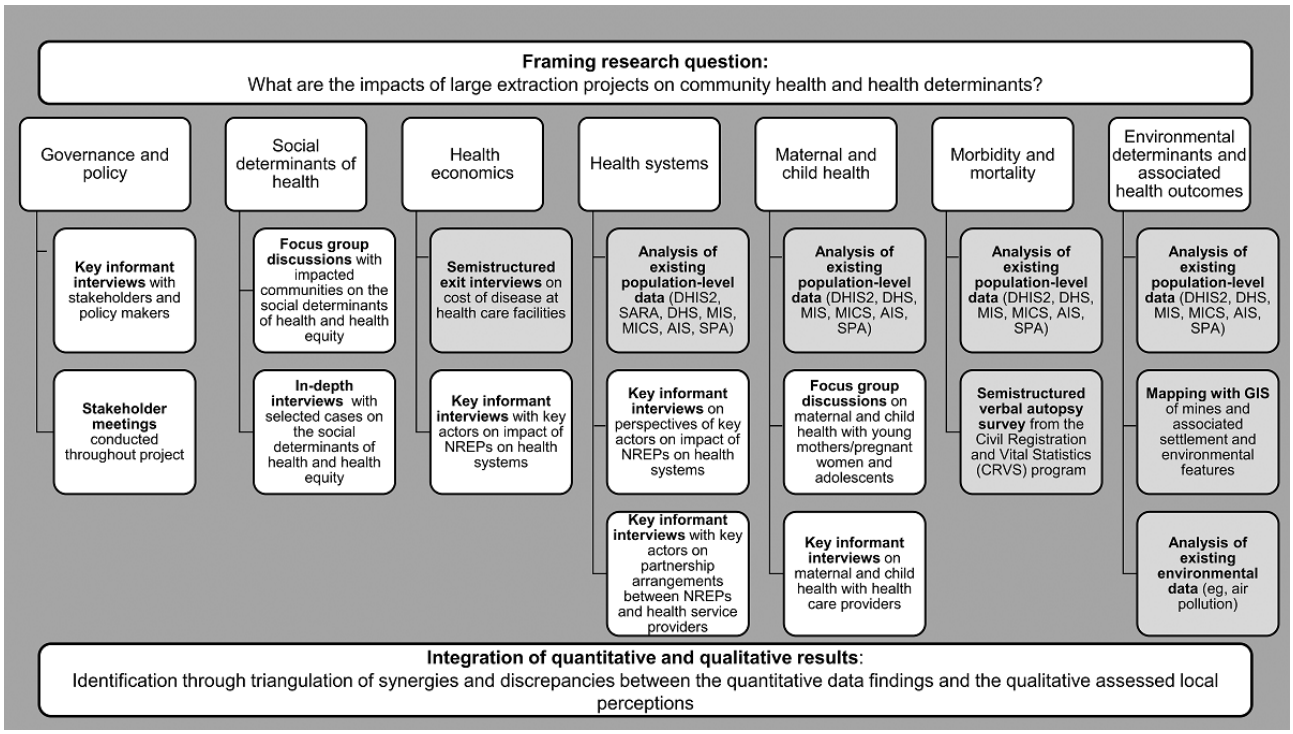
Methods

Summary

The study is designed as a concurrent triangulation mixed methods design [24], with simultaneous collection of (1) quantitative data used to measure resource extraction project effects on population health by describing incidence of disease and other health determinants and outcomes over time, as well as (2) qualitative data describing community perceptions of the impacts of projects on health. In this design, the data will be analyzed first separately and then integrated, with the advantage that each data source can complement and strengthen the findings of the other data sources, as well as drive further

research questions. Based on the initial study aims, seven major research topics of interest were identified: governance and policy, social determinants of health, health economics, health systems, maternal and child health, morbidity and mortality, and environmental determinants and associated health outcomes (see Figure 1). Data will be collected on every research topic by the local team within each country and then shared across all project teams, with the aim of making comparisons both within and between each country.

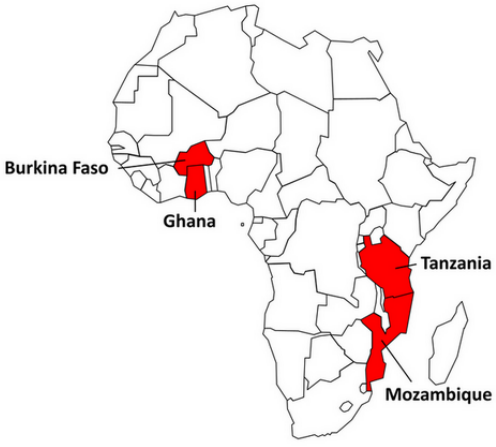
Figure 1. Study design. The headings are the research topics relevant to population health in natural resource extraction project (NREP) areas identified in conjunction with local partners. The specific research topics are investigated with both qualitative (white) and quantitative (gray) research methods. AIS: AIDS Indicator Survey; DHIS2: District Health Information System 2; DHS: Demographic and Health Survey; GIS: geographic information system; MICS: Multi-Indicator Cluster Survey; MIS: Malaria Indicator Survey; SARA: Service Availability and Readiness Assessment; SPA: Service Provision Assessment.



Setting

The health impacts of resource extraction projects vary based on the baseline characteristics and environment of the host community [25]. Therefore, we partnered with four different countries with a history of natural resource extraction and low health-related SDG index values: Burkina Faso, Ghana, Mozambique, and Tanzania (see Figure 2). Within those countries, three to four large mining sites were chosen for qualitative and quantitative field work. The mines were chosen based on type, size, length of operation, and type of ownership. Quantitative data on disease and health outcomes is available country-wide through District Health Information System 2 (DHIS2) and Demographic and Health Survey (DHS) databases (see Table 1) [26].

Figure 2. Location of study countries in sub-Saharan Africa.



In Burkina Faso, primary data collection was conducted in three gold mining sites: (1) Houndé Mine in Houndé district (population in 2006: 77,000); (2) Yaramoko Gold Mine in Bagassi district (population in 2006: 33,000); and (3) Bissa Mine in Kongoussi district (population in 2006: 71,000). The three mines have been operational since 2017, 2016, and 2013, respectively. While the Houndé and Bissa mines are open-pit mines, operations in Bagassi are underground. Houndé is a small city located approximately 5 km from the mine with formal and informal settlements reaching closer. In Bagassi, villages are scattered around the mining premises. Sabcé is the closest town from the Bissa mine. During the field visit, artisanal mining sites were observed around all three mines.

In Ghana, primary data collection was conducted in three gold and manganese mining sites: (1) Edikan Gold Mine (Perseus Mining) in the Upper Denkyira West district (population in 2019: 71,425); (2) Tarkwa (Anglo Gold Ashanti, Gold Fields, and Ghana Manganese Company) in the Tarkwa-Nsuaem Municipal district (population in 2019: 117,550); and (3) Newmont Ahafo Mine (Newmont Goldcorp) in the Asutifi North Municipal (population in 2010: 52,259). The three mines have been operational since 2011, 1961, and 2003, respectively. All three are open-pit mines. Edikan Gold Mine has four nearby communities—Ayanfuri, Fobinso, Nkonya, and Abenabena—located in both the Upper Denkyira West and Amenfi West districts. The sites near Tarkwa include the small communities of Akoon, Tarkwabanso, Wangarakrom and Badukrom, and Iduaprim. Four communities are near the Newmont Ahafo Mine: Tutuka, Kenyase 1, Kenyase 2, and Ntrotroso.

In Mozambique, primary data collection was conducted in three types of mining sites—ruby, titanium, and coal—involving communities of four administrative districts and four mining companies: (1) Montepuez Ruby Mining (MRM) in Montepuez district (population in 2017: 261,235); (2) Kenmare Moma Mining in Larde district (population in 2017: 85,971) and Moma district (population in 2017: 310,706); (3) Vale Mozambique; and (4) International Coal Ventures (ICVL), the latter two in Moatize district (population in 2017: 343,546). All four are open-pit mines that have been operating since 2007 (Kenmare Moma Mining) and 2011 (MRM, Vale, and ICVL). MRM is surrounded by four small villages—N´sewe, N´thorro, M´pene, and Namanhumbir—whose main economic activities are agriculture and unregulated artisanal mining. Near Kenmare Moma Mining, Moma and Larde are the main and small emerging coastal towns located approximately 80 km and 20 km from the Kenmare mining company, respectively. The company activities affect communities from Moma district (ie, Pilivili locality) and Larde district (ie, Topuito locality). The eight affected villages at Pilivili locality (ie, Moma district) are the closest, located 5 km away from the mining company. There are 11 neighborhoods and villages affected by Kenmare Moma Mining, of which seven are located within the mining concession area and four are less than 100 meters from the mining pit: Tipane, Mutiticoma, Izoua, and Topuito. Agriculture and fishing are the main economic activities; however, commerce is an activity emerging mainly after mining implementation in 2005. Moatize is a small town 20 km from the city of Tete. It is an industrial complex composed of at least six large-scale coal mining companies—Vale Mozambique, ICVL, Nkodezi Coal Company, Minas de Revubue, Minas de Moatize Riversdale Mozambique Limitada, and Jindal Mozambique Minerals (JSPL)—some being subsidiaries of Vale Mozambique. At the time of site visit, very active commerce activities were observed along the main road. Apart from the town of Moatize, there are 12 small affected communities surrounding mining companies, four belonging to Moatize-sede locality—Catete, Mphandue, Matambanhama, and Ntchenga—and eight to Benga locality—Cancope, Capanga Gulo, Campanga lowane, Chitambo, Chitondo, Nyambalualu, Kangale, and Benga-sede. The last five communities are along the Zambeze River, the main local source of water. Agriculture and fishing are the main economic activities.

In Tanzania, three gold mining sites were chosen for fieldwork: Geita Gold Mine (Geita district), Buzwagi Gold Mine (Kahama district), and Bulyanhulu Gold Mine (Mslala district). Geita and Buzwagi are open-pit mines, while Bulyanhulu is an underground mine (based on observation). At the time point of data collection (ie, March to May 2019), the Geita Gold Mine was fully operational and the Buzwagi and Bulyanhulu mines were both in reduced production status, meaning that they were processing already extracted material and no longer extracting new raw material. The Geita Gold Mine, about 70 km south of Lake Victoria, is located between several villages and Geita, a main city of the district. The Buzwagi mine is surrounded by several villages and is 6 km away from Kahama, the capital of the district. The vibrant villages of Kakola, Bushing'we, and Kakola Namba 9 are next to the Bulyanhulu mine. During the field visit, artisanal mining sites were observed around all three mines.

Table 1. Primary quantitative outcomes and data availability from the District Health Information System 2 (DHIS2), the Demographic and Health Survey (DHS), and other national-level databases.

Primary quantitative outcomes	Available in the DHIS2 at the district level	Available at the district level through the DHS or another data source
Health-related SDG[a] target indicator		
Stunting rate among children below the age of 5 years	Burkina Faso Ghana Mozambique Tanzania	Burkina Faso Tanzania
Maternal deaths per 100,000 live births	Burkina Faso Ghana Mozambique Tanzania	Burkina Faso Mozambique Tanzania
Proportion of births attended by skilled health personnel	Burkina Faso Ghana Mozambique Tanzania	Burkina Faso Mozambique
Under-5 mortality rate (deaths per 1000 live births)	Burkina Faso Ghana Mozambique	Burkina Faso Mozambique Tanzania
Number of new HIV infections per 1000 uninfected population members (by age group, sex, and key populations)	Burkina Faso Ghana Mozambique Tanzania	Mozambique
Tuberculosis incidence per 1000 persons per year	Burkina Faso Mozambique	Burkina Faso Mozambique Tanzania
Malaria incident cases per 1000 persons per year	Burkina Faso Ghana Mozambique	Mozambique Tanzania
Rates of noncommunicable diseases	Burkina Faso Ghana Mozambique	Mozambique Tanzania
Number of road traffic fatal injury deaths within 30 days, per 100,000 population members (age-standardized)	Burkina Faso	Burkina Faso Ghana Mozambique Tanzania
Health worker density and distribution	Burkina Faso	Burkina Faso Ghana Mozambique Tanzania
Additional health indicators for monitoring health for the SDGs and health system performance		
Number and distribution of health facilities per 10,000 population members	Burkina Faso Mozambique	Burkina Faso Ghana Mozambique Tanzania
Number of health workers per 10,000 population members	Burkina Faso	Burkina Faso Ghana Mozambique Tanzania

Primary quantitative outcomes	Available in the DHIS2 at the district level	Available at the district level through the DHS or another data source
Number of outpatient visits per 10,000 population members per year	Burkina Faso Mozambique	Burkina Faso Mozambique Tanzania
Vaccination coverage in children aged 12-23 months	Burkina Faso Ghana Mozambique Tanzania	Burkina Faso Ghana Mozambique Tanzania
Acute respiratory disease rate in children aged under 5 years	Burkina Faso Ghana Tanzania	Burkina Faso Ghana Mozambique Tanzania
Diarrhea rate in children aged under 5 years	Burkina Faso Ghana Mozambique Tanzania	Burkina Faso Ghana Mozambique Tanzania
Helminthic infection rate in different age groups	Burkina Faso Ghana	Burkina Faso Ghana Mozambique
Syphilis rates in children and pregnant women	Burkina Faso Ghana Mozambique	Burkina Faso Ghana Mozambique
Anemia rate in children aged under 5 years and pregnant women	Burkina Faso Ghana Mozambique Tanzania	Burkina Faso Ghana Mozambique Tanzania
Hypertension rate in adults	Burkina Faso Ghana Tanzania	Burkina Faso Ghana Mozambique Tanzania
Chronic respiratory tract infections rate among different age groups	Burkina Faso Tanzania	Burkina Faso Mozambique Tanzania

[a]SDG: Sustainable Development Goal.

Quantitative Study Components

The main quantitative component of the study will be a retrospective observational longitudinal study using routine health data extracted from DHIS2 and other national-level databases. Specifically, this quantitative part of the study seeks to answer the following questions: (1) Which districts have been directly, indirectly, or not impacted by projects? (2) What differences in health-related SDG indicators and other health indicators (eg, health systems) can be observed in districts impacted by resource extraction projects compared to nonimpacted districts, including maternal and child health–related indicators? (3) What are the costs and benefits to the health system of project implementation? (4) How do projects impact on environmental determinants of health? and (5) What are the strengths and limitations of the national, routine, health information systems and other data sources and repositories at the national level to monitor how projects impact on health-related SDG target indicators and other health indicators? The specific research topics belonging to this work package are effects on health systems, morbidity and mortality, environmental determinants of health, maternal and child health, and health economics. The research in this work package uses existing DHIS2 surveillance data routinely collected by the governments of the study countries, along with supplementary national datasets—DHS, Service Availability and Readiness Assessment (SARA), Malaria Indicator Survey (MIS), Multi-Indicator Cluster Survey (MICS), AIDS Indicator Survey (AIS), and Service Provision Assessment (SPA) (see Table 2)—and remote sensing data (eg, geographical positioning, weather, and pollution data).

Table 2. Sources of existing national-level data used in the HIA4SD (Health impact assessment for sustainable development) study.

Data source	Abbreviation	Description	Data frequency
District Health Information System 2	DHIS2	The DHIS2 is a Web-based, open-source, health information management and visualization system used in more than 40 countries, including governmental agencies in the project countries. The DHIS2 databases contain information about many of the health-related Sustainable Development Goals (SDGs), as well as other key health indicators for monitoring health and health system performance.	Monthly
Demographic and Health Survey	DHS	The DHS is a nationally representative household survey that provides information on a wide range of health, population, and nutrition indicators.	Usually every 5 years
Service Availability and Readiness Assessment	SARA	The SARA is a health facility assessment survey designed to monitor service readiness and availability indicators.	Every 2-3 years (country dependent)
Malaria Indicator Survey	MIS	The MIS is a household survey designed to collect information on a wide range of malaria indicators.	Every 3-5 years (country dependent)
Multi-Indicator Cluster Survey	MICS	The MICS is a household survey designed to collect data on indicators related to the situation of children and women.	Every 5 years (country dependent)
AIDS Indicator Survey	AIS	The AIS is a household survey designed to collect information on indicators related to HIV/AIDS.	Every 5-6 years (country dependent)
Service Provision Assessment	SPA	The SPA is a health facility assessment survey designed to evaluate health service delivery.	Every 9-10 years (country dependent)

Data Analysis

Geographic Information System Mapping Data

As part of the project, information on natural resource extraction projects (ie, type, size, and associated projects) and their exact geographic position will be extracted from government and other public record databases in the four study countries and mapped using ArcGIS software (Esri), a geographic information system (GIS). The data exist in diverse sources in different countries (ie, Ministries of Land and Natural Resources, Ministries of Water and Sanitation, Ministries of Science and Environment, Ministry of Energy, Forest Commissions, and local government records) as well as international monitoring organizations (ie, the International Finance Corporation [IFC] and the World Bank). Satellite imagery from the Landsat database will be used to estimate the impact of mining on settlement growth. The extracted satellite scenes will cover a time range spanning the years prior and after the opening of selected large-scale mining projects. With the aid of ancillary ground-truth information from historic Google Earth imagery, the size of settlements will be determined annually using machine learning algorithms. The annual growth of settlements will then be compared between mining and comparison sites.

Sampling of Natural Resource Extraction Project–Impacted Districts and Matching Comparison Districts

To analyze the differences in health outcomes between districts with and without extractive industries, a matched geographical analysis will be performed within the framework of the HIA4SD project to provide a stratified sampling framework. In a first stage, natural resource extraction projects will be mapped in each country together with their key attributes (ie, size and number of workers, length and age, type, and geographical footprint of the project). Each country will be spatially stratified into four levels for sampling purposes: (1) areas in direct proximity to a project (*highly impacted* areas), (2) areas within a 20-30 km buffer zone of the directly impacted regions (*impacted* areas), (3) regions bordering project areas that contain access roads or other economic or physical links to the project (*potentially impacted* areas), and (4) regions greater than 30 km away from a project that do not contain access roads or other economic or physical links to the project (*nonimpacted* areas). Comparison study sites will be selected from the *nonimpacted* areas and matched with important baseline characteristics (ie, community socioeconomic activities, vegetation, altitude, and ecological zone) to the *highly impacted* areas. All facilities, including public and private, that fit within perimeter boundaries and are registered in the DHIS2 database will be selected. In a second stage, *highly impacted* districts will be matched to two or three *nonimpacted* comparison districts within each country. Districts will be matched based on important baseline characteristics (ie, population, urbanization and aggregate night satellite brightness, square area, number of health care facilities, and disease rates) during the year before project implementation in the *highly impacted* district.

Quality Assessment of District Health Information System 2 Data and Association of Health Indicators With Natural Resource Extraction Projects

In order to assess the quality and completeness of the DHIS2 data, a comparison of DHIS2 and other data repositories with health statistics being collected at the local level under other work packages will be done. Next, potential positive and negative associations between health outcomes and health determinants (independent variables) and the existence of natural resource extraction projects (dependent variable) will be analyzed in the fourth working step by means of quasi-Poisson regression models and binomial regression models. The time span to be analyzed will be determined by the development history of the projects of interest in combination with the

availability of data, which will vary between datasets. For the regression model, setting-specific cluster effects (eg, urban, rural, and type of project) will be taken into account. In order to maximize statistical power, pooled cross-country analysis and meta-analysis will be employed in addition to country-specific evaluations. Finally, based on the previous working steps, strengths and limitations of the national routine health information systems and other datasets at the national level to monitor impacts on health-related SDG target indicators and other health indicators of extraction projects will be determined and systematically reported [27].

Economic Cost-Benefit Analysis

Costs to the health system and to the households arising from prespecified disease conditions in each country will be estimated based on a retrospective approach. Health system costs will be collected from published information and on-site in selected health facilities in impacted districts through key informant interviews. Cost-generating components, such as required medical resources related to the corresponding illness, will be identified and a monetary value will be attributed to them. Costs incurred by the households, which are associated with the health care received, will be obtained through exit interviews in health facilities. The combination of excess cases associated with the presence of resource extraction projects (ie, data generated in the quantitative part of the study) and corresponding costs in each study country will allow the estimation of the cost incurred to the health system and households. A cost-of-illness analysis will be employed for estimating the costs incurred because of specific diseases or conditions (eg, HIV incidence rates, accidents, and chronic respiratory diseases) that have been identified in the first work package as being significantly increased due to the presence of extractive industry projects. The comparison of incidence data from *impacted districts* with incidence data from matching *comparison districts* (ie, districts with similar characteristics as *impacted districts* but without the presence of extractive industries) will allow for calculating the number of excess cases (eg, per 100,000 inhabitants) over the duration of 1 year. Probabilistic sensitivity analysis will be employed to allow for uncertainty around the cost estimates. Economic benefits of resource extraction projects will be measured monetarily in terms of direct and indirect financial contributions from the projects to the health sector. Financing of health infrastructure is considered to be a direct financial contribution, while the share of the health budget in the incremental tax revenues is considered to be an indirect contribution. We will also measure any other contributions.

Qualitative Data Analysis

The key informant interviews, focus group discussions, and in-depth interviews will be recorded using digital voice recorders for subsequent transcription. The transcripts will then be imported into software for qualitative data analysis—NVivo (QSR International)—to code the transcripts for thematic content and framework analysis based on the COREQ (COnsolidated criteria for REporting Qualitative research) criteria [28,29]. The information obtained from different sources will be used for systematically assessing the different topics of the PhD projects (ie, partnership arrangements and the perception of health care services availability and accessibility), with a particular angle on maternal and child health, sexual and reproductive health of adolescent girls, and key social determinants linked to the perceived health impacts.

Triangulation of Quantitative and Qualitative Results

Using triangulation methodologies, the most striking quantitative and qualitative results will be compared within and across the research topics and countries. Synergies and discrepancies in national data sources and local perspectives will be identified and used to develop new research questions and tools. In addition, the qualitative results will be used to explore whether or not the national-level data adequately capture the full range of health impacts from resource extraction projects as reported by the local community.

Results

Fieldwork for the study began in February 2019 and concluded in February 2020. At the time of submission, qualitative data collection had been completed in Burkina Faso and Tanzania and is ongoing in Mozambique and Ghana. In Burkina Faso, 36 focus group discussions and 74 key informant interviews were conducted in three sites. In Ghana, 34 focus group discussions and 64 key informant interviews were conducted in three sites. In Mozambique, 75 focus group discussions and 103 key informant interviews were conducted in four sites. In Tanzania, 36 focus group discussions and 84 key informant interviews were conducted in three sites. Quantitative data extraction and collection is ongoing in all four study countries. Ethical approval for the study was received in all four study countries prior to beginning the fieldwork. Data analyses are underway and results are expected to be published in 2020 and 2021.

Discussion

The interlinkages between health, environment, social equity, politics, and economy are extraordinarily complex in regions dominated by extractive industries. To investigate these intersectoral dependencies, we have designed an innovative new study that is both mixed methods and multi-focus, allowing for a flexible research design to address many specific research questions under the umbrella of HIA. By utilizing a mixed methods design that triangulates both quantitative and qualitative findings, the study offers the opportunity to look at the impacts of natural resource extraction projects on population health from many different approaches and vantage points, allowing for a more complete understanding of the social, institutional, and economic changes that follow implementation of a new extractive project. This type of research cannot exist in an academic vacuum; to ensure sustainable policy development we have also incorporated the engagement of key stakeholders in every phase of the study with the aid of research on national and international governance, with active participation at all levels of the host country from policy and political leaders at the national level to everyday community members affected by resource extraction projects. Long after the study has concluded, a new resource of HIA will remain in-country in the form of

highly trained PhD candidates identified by the partner institutes that are already active in the health research of their countries.

The study has many strengths stemming from its novel design, in particular the complementary nature of the data collection due to the mixed methods design. By integrating many data sources and types, including national-level datasets, household data collected by international bodies, and primary data collected by diverse qualitative and quantitative methodologies, the study strengthens its evidence base while avoiding relying too much on one data source that may be biased or characterized by missing data. In particular, the in-depth multi-focus qualitative research across the four study countries means that the study captures a complete picture of health "on the ground" in diverse regions affected by natural resource extraction, instead of relying only on national-level data sources. The participatory and flexible approach to the study design means that the specific research topics and questions were largely driven by the ideas of researchers in the partner countries themselves, as opposed to a traditional top-down approach. This design feature ensures that findings will be relevant to policy makers at the national level, instead of mirroring priorities of the more distant international research community. Finally, instead of a siloed research approach where all researchers work separately to gather their own data and analyze it, data will be generated and shared across all research countries and topics, enabling within- and between-country comparisons. This multidisciplinary research will generate new hypotheses and findings across a wide range of health concerns, while promoting HIA methodology as an underutilized resource.

The DHIS2 database deserves special mention here as a new solution to data management across sub-Saharan Africa that is still relatively underutilized, despite being widely available for the past 5 years. By extracting and analyzing a wide range of health indicators countrywide for the past 5 years in all study countries, our study aims to realize the potential of DHIS2 in opening a new era of data being generated and analyzed in-country. This aim has particular reverberations for enabling countries using DHIS2 to start tracking their own progress toward the SDGs, instead of relying on estimates from international bodies.

There are some potential limitations to our study. The DHIS2 database is a relatively new data source, and little is known about the quality of its data. In addition, DHIS2 represents individuals that sought medical care from health care facilities and is likely underreporting actual incidence of disease, especially in areas without a strong reporting system or for less-severe disease that is not always treated in formal health care facilities. For some resource extraction projects, routine DHIS2 data may not be available before project implementation, making it difficult to approximate a true baseline health level in the impacted district. However, trends over time can still be described. In addition, nonimpacted districts that are matched as closely as possible on key demographic characteristics to impacted districts will provide a comparison that avoids the need for a baseline. These limitations are outweighed by the potential benefit of a comprehensive longitudinal assessment of health and extractive industries for the first time. In addition, as part of the study a quality assurance analysis will be done to evaluate the strengths and weaknesses of this data, which will aid researchers to better understand any biases or missing data that may affect the findings and also potentially identify areas for improvement in public health surveillance in the study countries. Finally, the primary quantitative data collection will be cross-sectional and, therefore, inferences about causality will be limited, although this will remain an important insight in an area where little previous research has been done. The qualitative research will be, by necessity, of a nonrandom sample of impacted community members, and the use of a "gatekeeper" for recruitment may result in a nonrepresentative sample; however, by recruiting across many communities and countries and using multiple gatekeepers where possible, the potential for bias will be limited as much as possible. By including a diversity of data sources in our study, many of these limitations will be reduced or eliminated.

The HIA4SD study is an innovative study designed to give a more complete picture, to our knowledge than ever before, of population health in regions affected by extractive industry activity. This will allow host countries to move closer to sustainable development by harnessing the potential positive socioeconomic impacts of these projects while minimizing negative consequences due to migration, disease, and environmental changes. By taking a thematic approach that incorporates a multifaceted mixed methods research design, using both already-existing secondary data to track change in project-impacted regions over time and primary data collected in the field, the project is able to capture the perspective of the many actors that have a stake in sustainable development in these regions while reducing project reliance on any one data source that may be incomplete. Finally, the integration of research on governance systems and engagement with key stakeholders into the project from the beginning helps situate and translate our research findings into concrete policy recommendations, aiding progress toward the SDGs.

Acknowledgments

We are most grateful to the Swiss Programme for Research on Global Issues for Development (r4d programme) for supporting the HIA4SD project. This work is part of the r4d programme [30], which is a joint funding initiative by the Swiss Agency for Development and Cooperation (SDC) and the Swiss National Science Foundation (SNSF) (grant number 169461).

Authors' Contributions

AF drafted the manuscript, with input from all coauthors. MSW is the principal investigator of the project and is responsible for overall study coordination. FB is a coprincipal investigator. HC, BN, IL, and HRZ are responsible for coordinating fieldwork and research in their respective countries.

Conflicts of Interest

None declared.

References

1. Listorti JA. Bridging Environmental Health Gaps: Lessons for Sub-Saharan Africa Infrastructure Projects. Main Report (Volume I). Washington, DC: The World Bank, Africa Technical Department, Environmentally Sustainable Development Division; 1996 May. URL: http://documents.worldbank.org/curated/en/963121468740127395/pdf/multi-page.pdf [accessed 2020-03-31]
2. Langston JD, Lubis MI, Sayer JA, Margules C, Boedhihartono AK, Dirks PH. Comparative development benefits from small and large scale mines in North Sulawesi, Indonesia. Extr Ind Soc 2015 Aug;2(3):434-444. [doi: 10.1016/j.exis.2015.02.007]
3. Knoblauch A, Hodges M, Bah M, Kamara H, Kargbo A, Paye J, et al. Changing patterns of health in communities impacted by a bioenergy project in northern Sierra Leone. Int J Environ Res Public Health 2014 Dec;11(12):12997-13016 [FREE Full text] [doi: 10.3390/ijerph111212997] [Medline: 25514152]
4. Afrane Y, Zhou G, Lawson B, Githeko A, Yan G. Effects of microclimatic changes caused by deforestation on the survivorship and reproductive fitness of Anopheles gambiae in western Kenya highlands. Am J Trop Med Hyg 2006 May;74(5):772-778. [Medline: 16687679]
5. Shandro J, Koehoorn M, Scoble M, Ostry A, Gibson N, Veiga M. Mental health, cardiovascular disease and declining economies in British Columbia mining communities. Minerals 2011 Oct 28;1(1):30-48 [FREE Full text] [doi: 10.3390/min1010030]
6. Reeson A, Measham T, Hosking K. Mining activity, income inequality and gender in regional Australia. Aust J Agric Resour Econ 2012 Jan 20;56:302-313 [FREE Full text] [doi: 10.1111/j.1467-8489.2012.00578.x]
7. von der Goltz J, Barnwal P. Department of Economics, Columbia University. 2014 Feb 23. Mines: The local welfare effects of mineral mining in developing countries URL: https://academiccommons.columbia.edu/doi/10.7916/D88D04VH/download [accessed 2019-07-23]
8. Campbell C. Migrancy, masculine identities and AIDS: The psychosocial context of HIV transmission on the South African gold mines. Soc Sci Med 1997 Jul;45(2):273-281. [doi: 10.1016/s0277-9536(96)00343-7]
9. Hilson G. An overview of land use conflicts in mining communities. Land use policy 2002 Jan;19(1):65-73. [doi: 10.1016/s0264-8377(01)00043-6]
10. Westwood E, Orenstein M. Does resource development increase community sexually transmitted infections? An environmental scan. Extr Ind Soc 2016 Jan;3(1):240-248. [doi: 10.1016/j.exis.2015.10.008]
11. O'Faircheallaigh C. Social equity and large mining projects: Voluntary industry initiatives, public regulation and community development agreements. J Bus Ethics 2014 Aug 8;132(1):91-103. [doi: 10.1007/s10551-014-2308-3]
12. Gibb H, O'Leary KG. Mercury exposure and health impacts among individuals in the artisanal and small-scale gold mining community: A comprehensive review. Environ Health Perspect 2014 Jul;122(7):667-672 [FREE Full text] [doi: 10.1289/ehp.1307864] [Medline: 24682486]
13. Molina-Villalba I, Lacasaña M, Rodríguez-Barranco M, Hernández AF, Gonzalez-Alzaga B, Aguilar-Garduño C, et al. Biomonitoring of arsenic, cadmium, lead, manganese and mercury in urine and hair of children living near mining and industrial areas. Chemosphere 2015 May;124:83-91. [doi: 10.1016/j.chemosphere.2014.11.016] [Medline: 25434277]
14. Yabe J, Nakayama SM, Ikenaka Y, Yohannes YB, Bortey-Sam N, Oroszlany B, et al. Lead poisoning in children from townships in the vicinity of a lead-zinc mine in Kabwe, Zambia. Chemosphere 2015 Jan;119:941-947. [doi: 10.1016/j.chemosphere.2014.09.028] [Medline: 25303652]
15. Hossain D, Gorman D, Chapelle B, Mann W, Saal R, Penton G. Impact of the mining industry on the mental health of landholders and rural communities in southwest Queensland. Australas Psychiatry 2013 Mar;21(1):32-37. [doi: 10.1177/1039856212460287] [Medline: 23344802]
16. Viliani F, Edelstein M, Buckley E, Llamas A, Dar O. Mining and emerging infectious diseases: Results of the Infectious Disease Risk Assessment and Management (IDRAM) initiative pilot. Extr Ind Soc 2017 Apr;4(2):251-259. [doi: 10.1016/j.exis.2016.08.009]
17. Quigley R, den Broeder L, Furu P, Bond A, Cave B, Bos R. Health Impact Assessment: International Best Practice Principles. Special Publication Series No. 5. Fargo, ND: International Association for Impact Assessment (IAIA); 2006 Sep. URL: http://hiaconnect.edu.au/old/files/HIA_Best_Practice_Principles.pdf [accessed 2019-07-24]
18. Harris-Roxas B, Viliani F, Bond A, Cave B, Divall M, Furu P, et al. Health impact assessment: The state of the art. Impact Assess Proj Apprais 2012 Mar;30(1):43-52 [FREE Full text] [doi: 10.1080/14615517.2012.666035]

19. Burton I, Wilson J, Munn RE. Environmental impact assessment: National approaches and international needs. Environ Monit Assess 1983 Jun;3(2):133-150. [doi: 10.1007/bf00398843]
20. Morgan RK. Environmental impact assessment: The state of the art. Impact Assess Proj Apprais 2012 Mar;30(1):5-14. [doi: 10.1080/14615517.2012.661557]
21. Winkler MS, Krieger GR, Divall MJ, Cissé G, Wielga M, Singer BH, et al. Untapped potential of health impact assessment. Bull World Health Organ 2013 Jan 31;91(4):298-305. [doi: 10.2471/blt.12.112318]
22. Harris P, Viliani F, Spickett J. Assessing health impacts within environmental impact assessments: An opportunity for public health globally which must not remain missed. Int J Environ Res Public Health 2015 Jan 20;12(1):1044-1049 [FREE Full text] [doi: 10.3390/ijerph120101044] [Medline: 25608592]
23. Winkler MS, Adongo PB, Binka F, Brugger F, Diagbouga S, Macete E, et al. Health impact assessment for promoting sustainable development: The HIA4SD project. Impact Assess Proj Apprais 2019 Nov 28:1-8. [doi: 10.1080/14615517.2019.1694783]
24. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. 3rd edition. Thousand Oaks, CA: SAGE Publications; 2018.
25. Winkler MS, Krieger GR, Divall MJ, Singer BH, Utzinger J. Health impact assessment of industrial development projects: A spatio-temporal visualization. Geospat Health 2012 May;6(2):299-301. [doi: 10.4081/gh.2012.148] [Medline: 22639132]
26. Youssef D, Khoury C, Allouch G, Haydar K, Jouny A, Zreik H, et al. Use of District Health Information System (DHIS-2) for real time surveillance: Lebanon 2017. In: Proceedings of the EMPHNET 6th Regional Conference 2018. 2018 Mar 29 Presented at: EMPHNET 6th Regional Conference 2018; March 27-29, 2018; Amman, Jordan. [doi: 10.2196/10547]
27. Farnham A, Utzinger J, Kulinkina AV, Winkler MS. Using district health information to monitor sustainable development. Bull World Health Organ 2019 Nov 29;98(1):69-71. [doi: 10.2471/blt.19.239970]
28. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
29. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Oct 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
30. Swiss Programme for Research on Global Issues for Development (r4d programme). URL: http://www.r4d.ch/ [accessed 2020-03-31]

Abbreviations

AIS: AIDS Indicator Survey
COREQ: COnsolidated criteria for REporting Qualitative research
DHIS2: District Health Information System 2
DHS: Demographic and Health Survey
GIS: geographic information system
HIA: health impact assessment
HIA4SD: Health impact assessment for sustainable development
ICVL: International Coal Ventures
IFC: International Finance Corporation
JSPL: Jindal Mozambique Minerals
MICS: Multi-Indicator Cluster Survey
MIS: Malaria Indicator Survey
MRM: Montepuez Ruby Mining
r4d programme: Swiss Programme for Research on Global Issues for Development
SARA: Service Availability and Readiness Assessment
SDC: Swiss Agency for Development and Cooperation
SDG: Sustainable Development Goal
SNSF: Swiss National Science Foundation
SPA: Service Provision Assessment

Edited by G Eysenbach; submitted 20.11.19; peer-reviewed by DJ Kishore, T Ingviya, T Isaacs; comments to author 06.01.20; revised version received 03.02.20; accepted 22.02.20; published 08.04.20.

Please cite as:
Farnham A, Cossa H, Dietler D, Engebretsen R, Leuenberger A, Lyatuu I, Nimako B, Zabre HR, Brugger F, Winkler MS
Investigating Health Impacts of Natural Resource Extraction Projects in Burkina Faso, Ghana, Mozambique, and Tanzania: Protocol for a Mixed Methods Study
JMIR Res Protoc 2020;9(4):e17138
URL: http://www.researchprotocols.org/2020/4/e17138/
doi:10.2196/17138
PMID:32266876

Protocol

EXamining ouTcomEs in chroNic Disease in the 45 and Up Study (the EXTEND45 Study): Protocol for an Australian Linked Cohort Study

Celine Foote[1,2], MBBS, PhD; Carinna Hockham[1,3], DPhil; Louisa Sukkar[1,4], MBBS; Anna Campain[1,3], PhD; Amy Kang[1,3], MBBS; Tamara Young[1,3], MBBS; Alan Cass[5], MD, PhD; Clara K Chow[1,6,7], MBBS, PhD; Elizabeth Comino[8], PhD; Martin Gallagher[1,2,3,9], MBBS, PhD; Stephen Jan[1,3,9], PhD; John Knight[1,3], MBBS; Bette Liu[10], MBBS, DPhil; Martin McNamara[11], PhD; David Peiris[1,3], MBBS, PhD; Carol Pollock[12,13], MD, PhD; David Sullivan[9,14], MD; Germaine Wong[4,15], MBBS, PhD; Sophia Zoungas[1,16], MBBS, PhD; Kris Rogers[1,17], PhD; Min Jun[1,3], PhD; Meg Jardine[1,2], MBBS, PhD

[1]The George Institute for Global Health, Sydney, Australia
[2]Concord Repatriation General Hospital, Sydney, Australia
[3]Faculty of Medicine, University of New South Wales, Sydney, Australia
[4]School of Public Health, University of Sydney, Sydney, Australia
[5]Menzies School of Health Research, Charles Darwin University, Darwin, Australia
[6]Westmead Applied Research Centre, Faculty of Medicine and Health, University of Sydney, Sydney, Australia
[7]Department of Cardiology, Westmead Hospital, Sydney, Australia
[8]Centre for Primary Health Care and Equity, University of New South Wales, Sydney, Australia
[9]Sydney Medical School, University of Sydney, Sydney, Australia
[10]School of Public Health and Community Medicine, University of New South Wales, Sydney, Australia
[11]The Sax Institute, Sydney, Australia
[12]Renal Division, Kolling Institute for Medical Research, Sydney, Australia
[13]University of Sydney, Sydney, Australia
[14]Department of Chemical Pathology, Royal Prince Alfred Hospital, Sydney, Australia
[15]Centre for Transplant and Renal Research, Westmead Hospital, Sydney, Australia
[16]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
[17]Graduate School of Health, University of Technology Sydney, Sydney, Australia

Corresponding Author:
Carinna Hockham, DPhil
The George Institute for Global Health
1 King Street
Sydney
Australia
Phone: 61 432964138
Email: chockham@georgeinstitute.org.au

Abstract

Background: Chronic kidney disease (CKD) and diabetes are the major causes of death and disability worldwide. They are associated with high health service utilization persisting over many years. Their slow progression and wide clinical variation make them eminently suitable for study in population-based cohorts. However, current understanding of their prevalence, incidence, and progression is largely based on studies conducted in clinical populations.

Objective: This study aims to establish a novel link between an existing population-based cohort (the 45 and Up Study) and routinely collected laboratory and administrative data to facilitate research across the full disease spectrum of CKD and diabetes.

Methods: In the EXTEND45 Study (EXamining OuTcomEs in chroNic Disease in the 45 and Up Study), baseline questionnaire responses of over 260,000 participants of the 45 and Up Study aged ≥45 years living in New South Wales (NSW), collected between January 2006 and December 2009, are linked to data from laboratory service providers as well as national- and state-based

administrative datasets via probabilistic linkage. Routinely collected data were obtained for participants who could be linked between January 2005 and July 2013. Laboratory data will enable the identification of early cases of chronic disease and the assessment of clinically relevant biochemical targets during the disease course. Health administrative datasets will allow for the examination of health service use, pharmacological management, and clinical outcomes.

Results: The study received ethics approval from the NSW Population and Health Services Research Ethics Committee in February 2014. Data linkage for 267,153 of the 45 and Up Study participants was completed in June 2016, with congruent linkage achieved for 265,086 (99.23%) individuals. To date, the CKD and diabetes cohorts have been identified (published elsewhere), and a diverse portfolio of research projects relating to disease burden, risk factors, health outcomes, and health service utilization is in development.

Conclusions: The EXTEND45 Study represents an unparalleled opportunity to perform extensive research into diseases of considerable public health and clinical importance. Strengths include the population-based nature of the cohort and the availability of longitudinal information on the complete disease pathway for affected individuals.

International Registered Report Identifier (IRRID): RR1-10.2196/15646

(JMIR Res Protoc 2020;9(4):e15646) doi:10.2196/15646

KEYWORDS

chronic kidney disease; diabetes mellitus; cardiovascular disease; data linkage; biomarkers

Introduction

Chronic diseases such as chronic kidney disease (CKD) and diabetes are major drivers of death [1], disability [2], and health care costs globally [3-5]. In Australia, 1.7 million people had indicators of CKD based on biomedical testing in 2011-2012 [6], whereas recent estimates for diabetes suggest that 1.2 million individuals were living with type 1 or 2 diabetes in 2017-2018 [7]. Adults aged 65 years and older bear the majority of this burden. Estimates vary according to study context and design, but most studies indicate at least a twofold increase in the prevalence of diabetes or CKD in people aged >65 years compared with their younger counterparts [8-11]. As the population ages, this problem is expected to further increase [12], with important consequences for already stretched health systems.

Diabetes and CKD are intrinsically linked [13,14], with diabetes being an important risk factor for CKD, and both independently associated with an increased risk of cardiovascular disease (CVD) [15]. The relative risk of CVD in adults with diabetes compared with those without ranges from 1 to 3 in men and 2 to 5 in women [16]. Similarly, individuals with impaired kidney function and increased urinary albumin excretion have a twofold to fourfold higher risk of developing CVD than those whose kidney function is normal [17]. In patients with established CVD, coexisting diabetes or CKD is associated with worse outcomes and a higher mortality rate [18-20]. Early detection and timely intervention are, therefore, paramount to the prevention and management of this trifecta of conditions.

Although there is a growing understanding of the prevalence and progression of diabetes, CKD, and their CVD-related complications, many studies to date have taken place in clinical populations recruited through specialist clinics and hospital settings. Evidence for effective management strategies, including clinical targets (eg, for glycated hemoglobin [HbA_{1c}] in diabetes) and different therapy options (eg, renin-angiotensin-aldosterone blockade in CKD), largely stems from randomized trials conducted in controlled clinical settings with narrowly selected patient populations. However, the slowly progressive nature of diabetes and CKD, coupled with their wide clinical variation, means that earlier or milder stages of disease are missed in hospital- or clinic-based cohorts. More population-based research into these diseases is needed to (1) gain detailed knowledge of the incidence and prevalence of diabetes and CKD in the community, including earlier disease stages, (2) identify novel determinants of disease progression, (3) better assess the broad-scale benefits and adverse effects of evidence-based therapies, and (4) evaluate health care utilization, costs, and outcomes across the full disease spectrum.

Direct assessment of the various facets of chronic disease burden and management in prospective cohorts is a major undertaking. To generate reliable, generalizable estimates, cohort studies must include a sufficiently large sample to capture most cases and be adequately distributed, both geographically and socioeconomically. In addition, rigorous follow-up of study participants is necessary, which is costly and often results in high loss to follow up. Routinely collected and administrative clinical data—collected as a by-product of patient care—offer an alternative method for acquiring detailed longitudinal information on health service use, disease burden, and clinical outcomes in a large proportion of the population.

The 45 and Up Study is a large-scale study established in 2006 that combines the merits of a prospective population-based cohort study and linked administrative health data [21]. Comprising more than 260,000 individuals aged ≥45 years living in New South Wales (NSW), the 45 and Up Study was designed to obtain information on healthy aging through questionnaires delivered to participants. In addition, participants consented to their questionnaire responses being linked to routinely collected health data for research.

The infrastructure generated through the 45 and Up Study offers the unique opportunity to examine diabetes, CKD, and their associated health outcomes in the community. However, to date, there has been no effort to utilize valuable data generated from routine laboratory testing to examine early disease states or assess the achievement of clinical targets in patients with CKD,

diabetes, or both. The EXTEND45 Study (EXamining ouTcomEs in chroNic Disease in the 45 and Up Study) builds upon the 45 and Up Study by establishing a novel link between the 45 and Up Study participants and their clinical data held by laboratory service providers. The EXTEND45 Study will concurrently link the 45 and Up Study participants to a range of other administrative datasets to obtain information across the full spectrum of disease and health service provision, spanning 10 years.

The mission of the EXTEND45 Study is to establish a rich data resource to examine CKD, diabetes, and their CVD-related complications. Initial research objectives of the EXTEND45 Study include the following:

1. To determine the prevalence and incidence of CKD in the 45 and Up Study cohort and define a seminal CKD cohort that can be used in future studies;
2. To determine the prevalence and incidence of diabetes in the 45 and Up Study cohort and define a seminal diabetes cohort that can be used in future studies;
3. To determine the prevalence and incidence of CKD in individuals identified as having diabetes in the 45 and Up Study cohort;
4. To identify risk factors for CKD and diabetes in the community;
5. To examine the real-world management of these diseases, including prescribing patterns, and identify evidence-practice gaps in the care of individuals with CKD, diabetes, or both, particularly in relation to CVD prevention; and
6. To evaluate the attainment of clinical targets in the community, identify risk factors for nonattainment, and assess associated health outcomes.

This paper provides a detailed description of the data sources, linkage methods, and governance structure of the EXTEND45 Study.

Methods

The Australian Health System

Australia has a universal health system comprising a multifaceted network of government (public) and private providers. Medicare is the universal public health insurance scheme, which is funded by the federal government through a combination of general tax revenue and a Medicare levy based on taxable income. It is administered by Services Australia (formerly the Department of Human Services) and provides free or subsidized access to 3 main areas of health care provision: (1) medical and health services outside of the public hospital setting (administered through the Medicare Benefits Schedule [MBS]), (2) prescription pharmaceuticals (administered through the Pharmaceutical Benefits Scheme [PBS]), and (3) public hospital services (funded jointly with, and managed by, state and territory governments). In addition, the federal government pays 75% of the fee for services and procedures for private patients in public or private hospitals and subsidizes private health insurance. The EXTEND45 Study includes administrative data collected by the federal government (ie, MBS and PBS) as well as state-specific data collections of the NSW government.

The 45 and Up Study

Details of the 45 and Up Study have been published previously [21]. Briefly, the 45 and Up Study comprises adults aged ≥45 years sampled from the general population of NSW, Australia's most populous state. Between January 2006 and December 2009, participants were randomly sampled from the Services Australia enrollment database, with individuals aged ≥80 years or living in remote areas oversampled by a factor of 2. Consenting participants self-completed baseline questionnaires in English, which included questions on demographic and socioeconomic characteristics, personal health behaviors, medical and surgical history, medications, and physical and psychological health [22]. Participants also consented to long-term follow-up, including linkage of their baseline responses to routinely collected health datasets. In the EXTEND45 Study, laboratory and administrative datasets have been linked at an individual level to the 45 and Up Study participants and their baseline questionnaire responses.

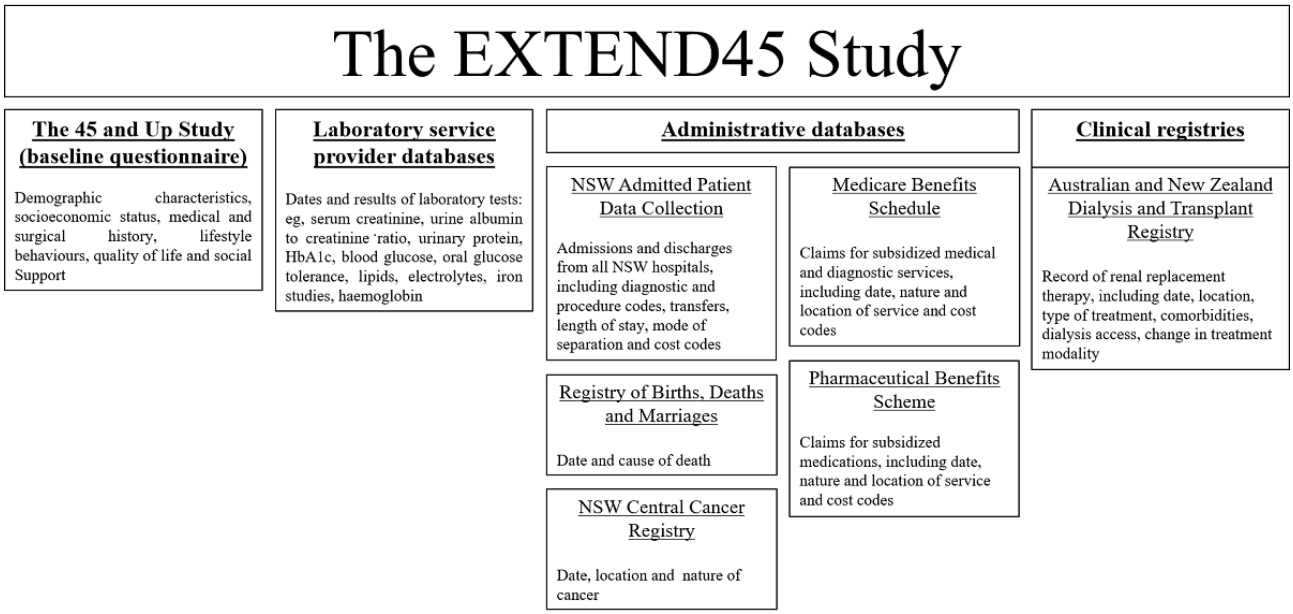
Linked Data Sources

Laboratory Service Providers

Government datasets hold information on referrals for community laboratory services but not the clinical results of those services. Instead, these are held by the laboratory service providers who conduct the tests, together with patient identifiers, tests performed, and date of testing (Figure 1). Multiple public and private laboratory service providers are active within NSW, and these vary in geographical scope and population coverage. In addition, providers offer varying levels of access to community, outpatient, or inpatient laboratory services, thereby capturing health information across different facets of an individual's health journey. Several major community laboratory service providers have been recruited for the EXTEND45 Study, and processes are underway to include data from an inpatient provider.

Available laboratory results and their meta-data between January 2005 and October 2015 are included at present. These enable early case identification (eg, serum creatinine measurements for kidney disease) as well as the assessment of disease progression, presence of disease complications (eg, albuminuria in patients with diabetes), and achievement of therapeutic targets (eg, HbA_{1C} in patients with diabetes and hemoglobin levels in patients with CKD).

Figure 1. An overview of the EXTEND45 Study (EXamining OuTcomEs in chroNic Disease in the 45 and Up Study) data sources and variables. NSW: New South Wales.



New South Wales Admitted Patient Data Collection

The NSW Admitted Patient Data Collection (NSW APDC) collates information on all admitted patient services provided by public (including psychiatric) and private hospitals, private day procedures centers, and public multipurpose services in NSW [23]. For public hospitals, data are recorded for each episode of care, with each episode defined as a period of stay in hospital before the patient is discharged or transferred, dies, or becomes a different *type* of patient (eg, acute, palliative, and rehabilitative). For private hospitals, each APDC record represents a complete hospital stay. Data include dates of admission and separation, referral source, diagnoses (including external causes), procedures, and service referred to on discharge. Up to 50 relevant diagnoses are included and are coded according to standardized codes from the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM).

At present, the EXTEND45 Study includes APDC data on separations that occurred between January 2005 and June 2014. NSW APDC data will be used to assess health service utilization, length of stay, referral and transfer patterns, and costs of hospital care (Figure 1).

Medicare Benefits Schedule and Pharmaceutical Benefits Scheme

Medicare operates by paying a specified benefit (in the form of a rebate) for services or prescription medicines that qualify for a benefit under the MBS and PBS, respectively, and for which a claim has been processed. As such, the MBS data collection contains information on all claims for medical and diagnostic services (including laboratory testing) provided to Australian citizens and permanent residents by registered medical and other practitioners, whereas the PBS data collection records all claims for prescription medicines above the PBS copayment threshold. From April 2012, suppliers of PBS medicines were required to provide data on the dispensing of medicines below the copayment threshold. MBS services and PBS items are coded using a system of item numbers listed in the relevant schedules.

MBS and PBS data are available from June 2004 to December 2016. MBS data will be used to assess health service utilization in primary care and outpatient settings, costs of care, and referral patterns. PBS data will be used to examine medication adherence, persistence, and cost, as well as the geographical distribution of services.

New South Wales Registry of Births, Deaths, and Marriages

All deaths are certified by a medical practitioner and registered by the NSW Registry of Births, Deaths, and Marriages (NSW RBDM). Details include cause and date of death, with cause of death coded using the ICD-10-AM coding system.

Data are available from February 2006 to March 2015. Access to NSW RBDM data will allow the comprehensive assessment of mortality, including primary and secondary causes of death.

Australian and New Zealand Dialysis and Transplant Registry

The Australian and New Zealand Dialysis and Transplant (ANZDATA) Registry maintains records of all patients in Australia and New Zealand with end-stage kidney disease who receive chronic renal replacement therapy (RRT), that is, dialysis or transplantation. The registry records the date of referral, start date of RRT, treatment modalities and vascular access for dialysis (both initial and any changes), and treatment outcomes for all patients.

Linkage to the ANZDATA Registry will capture information pertaining to the progression of kidney disease, including dialysis initiation or kidney transplantation. In addition, linkage of practice factors, such as treatment modalities and vascular access, to information on participant socioeconomic status and

health service use may help to identify novel areas for intervention (Figure 1).

New South Wales Central Cancer Registry

The NSW Central Cancer Registry (NSW CCR) is a registry of all patients with cancer in NSW. The data collected relates to invasive primary cancers and cancer deaths. It does not include skin cancers other than melanoma (eg, basal cell carcinoma and squamous cell carcinoma are excluded). Data collected by the NSW CCR includes clinical details describing the cancer, records of care from a notifier, pathology reports, and death certificates.

The inclusion of cancer registry data will allow further delineation of chronic disease groups at risk of cancer complications, with respect to comorbidities, lifestyle, socioeconomic status, and health service utilization.

Data Linkage Methods

Record linkage brings together information on the same individual from different data sources. Strict privacy protecting protocols are mandated by the NSW Population and Health Services Research Ethics Committee (PHSREC), the data custodians, and the NSW Center for Health Record Linkage (CHeReL). The CHeReL is an intermediate body established in 2006 to provide expertise in data linkage methodology and maintain a record linkage system that protects data privacy. It is jointly managed by the NSW Ministry of Health and the NSW Cancer Institute.

The Sax Institute has an ongoing link with Services Australia, and as a result has processes in place to deterministically link MBS and PBS data to their 45 and Up Study participants once a year. This is done using a unique identifier provided to the Sax Institute by Services Australia. For all other aforementioned datasets, linkage to 45 and Up Study data is performed by the CHeReL using probabilistic linkage (Figures 2 and 3). Data custodians provide the CHeReL with an encrypted unique identifier and relevant personal information (full name, date of birth, address, gender, and, where available, country of birth) for all patients over the relevant time-period. No clinical data are provided to the CHeReL at any point in the linkage process. The CHeReL uses personal information to link records from different data sources, using probabilistic linkage in the *ChoiceMaker* software (ChoiceMaker Technologies, Inc; Figure 2). For records with doubtful matches, a clerical review is also conducted [24]. Once records are linked, the corresponding individual is allocated a project-specific person number (PPN). The CHeReL sends the PPNs for all individuals to each data custodian, together with the original unique identifier used in that custodian's particular dataset (Figure 3). Each data custodian will now hold an EXTEND45-specific PPN for each individual in their respective dataset for whom linkage with at least one other dataset was achieved. The data custodian merges the PPN with the clinical variables that have been approved for use in the EXTEND45 Study and uploads the deidentified data to a highly secure server that is accessed by the EXTEND45-approved researchers. The PPN is used to combine records for the same person from different datasets.

Figure 2. A schematic diagram of the linkage process: Stage 1. Each data custodian provides the Centre for Health Record Linkage (CHeReL, a data linkage unit) with relevant personal information of individuals who have accessed their service, together with their unique record number. The names and details used in the figure are fictitious and do not relate to any participants of the 45 and Up Study. The CHeReL uses an algorithm to match individuals to participants of the 45 and Up Study and assigns a project-specific person number to each individual. DOB: date of birth.

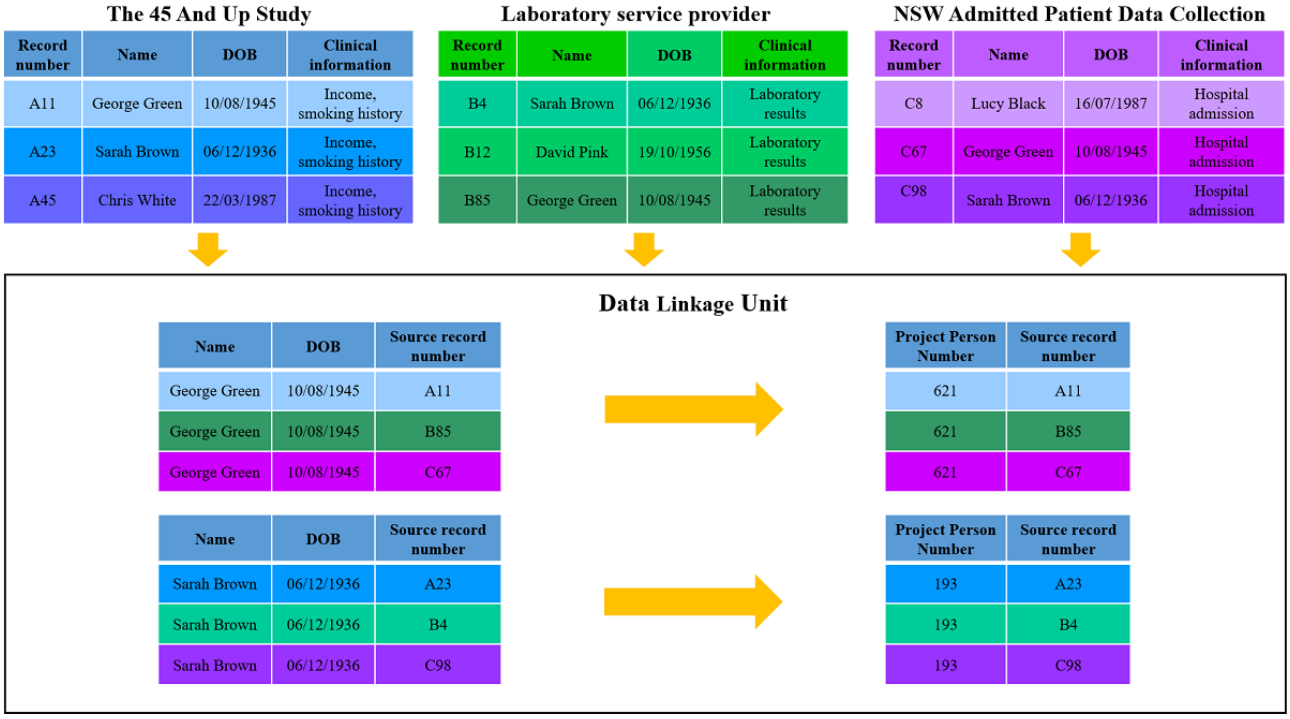
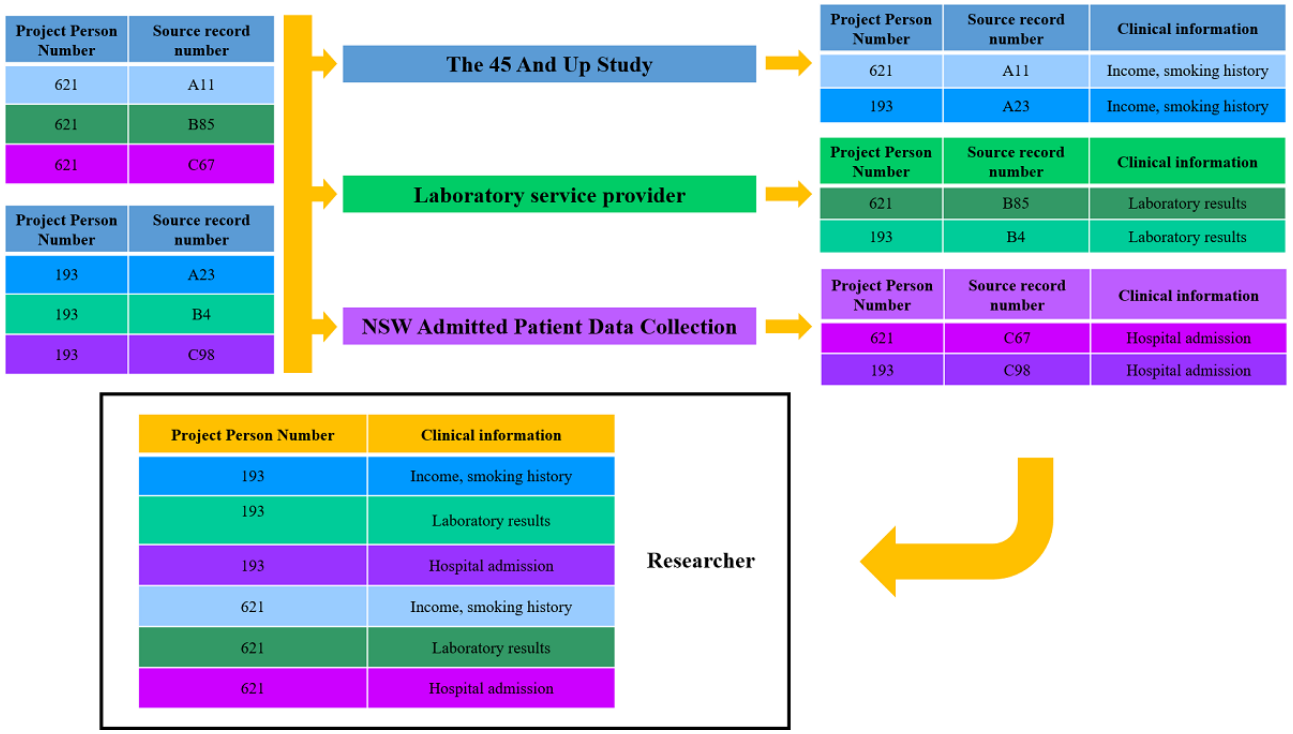


Figure 3. A schematic diagram of the linkage process: Stage 2. The Centre for Health Record Linkage (CHeReL) provides data custodians with the project-specific person number (PPN) for each individual, with all personal information removed. Data custodians use the PPN, together with their own record numbers, to extract the requested data and upload the data to Secure Unified Research Environment.



Ascertainment of Disease Status

The 2 initial diseases of interest are CKD and diabetes. For each individual, their diabetes and CKD status and, where appropriate, timing of incident disease, are determined using the various linked datasets. The precise methods used to ascertain disease status will be outlined in the relevant publications. In general, however, kidney function and CKD status are predominantly derived using data from laboratory records and, where appropriate, the ANZDATA Registry. Linked serum creatinine measurements are applied to the Chronic Kidney Disease Epidemiology Collaboration equation to calculate the estimated glomerular filtration rate [25]. Other available laboratory data related to kidney function include urine albumin-to-creatinine ratio and urine protein-to-creatinine ratio.

Diabetes status is derived using the following criteria: (1) recorded dispensing of insulin (Anatomical Therapeutic Chemical [ATC] Classification: A10A) or oral blood glucose-lowering medication (ATC Classification: A10B) in the PBS dataset, (2) self-reported diabetes on the 45 and Up Study (answered *Yes* to Q24. *Has a doctor EVER told you that you have diabetes?*), (3) laboratory record of HbA_{1c} result >6.5% on one occasion, (4) laboratory record of fasting plasma glucose levels >7.0 mmol/L, and (5) laboratory record of serum glucose levels >11.1 mmol/L. At present, no distinction has been made between type 1 and type 2 diabetes.

Study Governance

The EXTEND45 Study is led by a collaborative group of physicians, epidemiologists, and statisticians coordinated by The George Institute for Global Health, Sydney. Similar to the 45 and Up Study, the EXTEND45 Study has been designed to facilitate cooperative projects that lie within its scope, with oversight provided by a designated steering committee. The steering committee includes representatives from The George Institute for Global Health, the Universities of NSW and Sydney, the Sax Institute, clinical experts from Australian hospitals, and nonvoting representatives from pharmaceutical sponsors. Current steering committee members are listed in Multimedia Appendix 1.

The steering committee is responsible for the general oversight and governance of the study and provides academic independence and integrity. It provides scientific advice regarding all aspects of study design, conduct, analyses, and publication of results. In addition, the steering committee is responsible for the approval of prespecified analysis plans. Analysis proposals from internal or external researchers are encouraged and may be submitted by academic investigators, health service providers, or commercial health care entities.

Ethical Considerations and Data Privacy

Data collection for the EXTEND45 Study was through data linkage only. Participants are included on the basis of their participation and consent to the 45 and Up Study, which included consent for linkage to their routinely-collected health records. The 45 and Up Study was approved by the University of NSW Human Research Ethics Committee (UNSW HREC). Ethical approval for the EXTEND45 Study was obtained from the NSW Population and Health Service Research Ethics Committee (HREC/13/CIPHS/69).

Participant confidentiality is protected in several ways. First, linkage methodology splits the processes of record linkage and data analysis, ensuring that participant identifiers are always kept separate from clinical data. Second, security measures in place at the CHeReL ensure that the risks of a breach of privacy

are minimal. Third, as mandated by the Sax Institute, all data within the EXTEND45 Study is transferred and accessed via the Secure Unified Research Environment (SURE). SURE is a high-performance remote-access virtual computing environment designed specifically for secure access, storage, and analysis of anonymized health information [26]. To prevent data from being exchanged between different linked data studies, each study in the SURE facility exists within its own security perimeter.

Results

The 45 and Up Study cohort is relatively heterogeneous, with variation observed across most variables [27]. Of the 267,153 participants of the 45 and Up Study who completed the baseline questionnaire, data linkage was performed for 266,969 (99.93%). Incongruent dates, for example, between the 45 and Up Study enrollment and recorded date of death, or other erroneous dates were identified for 1883 individuals. The final linked dataset, therefore, comprises 265,086 individuals. As of November 2019, laboratory data are available for 152,169 (57.40%) individuals, with plans for expansion ongoing. Prevalent and incident cases of CKD and diabetes have been identified from multiple data sources, and the corresponding prevalence and incidence estimates have been presented in separate publications. A diverse portfolio of research questions relating to CKD and diabetes and their management in the community are currently underway.

Discussion

Principal Findings

Through a novel linkage between the 45 and Up Study, various administrative databases, and data from laboratory service providers, the EXTEND45 Study represents a unique and rich data resource that can be used to investigate a range of questions relating to chronic disease. The study's mission is to provide much-needed evidence of the epidemiology, burden, progression, and clinical management of chronic diseases such as CKD and diabetes in the general Australian population and across the full disease spectrum. Multiple projects that use this data resource are already completed or underway.

Altogether, the datasets included in the EXTEND45 Study provide a more complete overview of the care provided to patients with chronic disease than is currently available. The 45 and Up Study baseline questionnaire included questions on a range of personal health behaviors (eg, smoking, alcohol consumption, physical activity, and sleep habits) as well as demographic, social, and economic characteristics (eg, marital status, country of birth, and education level). Although many of these factors are either known or hypothesized to be important confounders of chronic disease incidence and progression, they are typically poorly captured in administrative datasets. Conversely, the 45 and Up Study baseline data provide only a single snapshot in time, are exclusively self-reported, and are lacking in granularity. In the EXTEND45 Study, the combining of NSW APDC and other health datasets will allow outcomes, comorbidities, and complications of CKD and diabetes to be examined. Moreover, laboratory data are critical for identifying individuals at the early stages of their disease and allow more granular assessment of the achievement (or lack thereof) of clinical targets. Finally, the collection of postcodes of residences and health care facilities in the various datasets will enable the geographic evaluation of access and use of all tiers of health services.

The effective use of routinely collected data and, in particular, laboratory data in chronic disease research is demonstrated by the many research and policy outputs of the ongoing Alberta Kidney Disease Network (AKDN). The AKDN is a collaborative research organization that holds a central repository of linked laboratory and administrative data from Alberta, Canada [28]. Research using this repository has ranged from a comparative risk assessment of coronary events in patients with diabetes and CKD [20] to factors associated with RRT initiation [29]. In the United Kingdom, the UK Biobank, established in 2006-2010, has recruited 500,000 people aged 40 to 69 years to provide detailed information about themselves, undergo various measurements, provide biological samples for analysis, and have their electronic health record data linked, to create a major data resource. To date, 989 papers have been published that used this resource. Similarly, the UK Clinical Practice Research Datalink (CPRD) links general practice records to secondary care, mortality, and other disease-specific databases. The CPRD is an excellent example of the utility of real-world data in generating evidence for drug safety guidance and clinical practice in particular. Research carried out as part of the EXTEND45 Study will complement these earlier studies.

Strengths and Limitations

A main strength of the EXTEND45 Study is the breadth and diversity of the data held within it, allowing analyses into a range of research questions to be conducted. Some of the potential analyses have already been described. Importantly, the inclusion of laboratory data allows researchers to use the EXTEND45 Study dataset to identify earlier stages of both diabetes and CKD and move away from reliance upon self-report for diabetes. This will provide essential information on the prevalence and progression of early disease in the community as well as health care delivery to those affected with mild disease. Indeed, previous studies have either relied on clinical populations or used sources that only include outcome measures such as hospitalization or mortality—in both instances, chronic disease is identified later in its disease course. Second, multiple data sources will be used to ascertain chronic disease status, which will improve case identification. The relative contributions of different data sources to the identification of cases will be reported in the respective publications. Finally, the combined infrastructure of the 45 and Up Study and EXTEND45 Study offers a cheaper and more efficient alternative to traditional longitudinal cohort studies for assessing the incidence of disease, health outcomes, and health service utilization in a large cohort.

There are important limitations that must be considered when using the EXTEND45 Study database, both at the design stages and during interpretation of results. Routinely collected and administrative data are primarily collected for purposes other than research. As a result, not all variables of interest will be available all of the time. For example, within the EXTEND45 Study, there is no information on blood pressure, an important

clinical variable in both CKD and diabetes. Depending on the specific research question, surrogates for some of these missing variables may be deployed. For instance, for high blood pressure, a pharmaceutical dispensing of antihypertensive medication or hospital admission for hypertension-related problems may be used. A second limitation is that CKD is primarily defined using laboratory tests. As a result, disease identification is limited to individuals who have presented for medical review and had laboratory tests performed. Conversely, the *absence* of a linked laboratory test in the EXTEND45 Study dataset may be due to (1) the test being performed by a laboratory service provider not part of the EXTEND45 Study, (2) an individual having no indication for a test (and, therefore, healthier), or (3) an individual being indicated but not having access to the appropriate health services. Given the countervailing effects of these different possibilities, a comparison of individuals with linked laboratory data to those without linked laboratory data must be performed in individual analyses. The use of routine laboratory results over time means that inter- and intralaboratory variation may exist. The effects of this are minimized through deep cleaning of the data to identify potentially anomalous results or inconsistent measurement units as well as detailed discussions with laboratory service providers to construct a data dictionary that is specific to their dataset.

Finally, the response rate for the 45 and Up Study baseline questionnaire was 18%, which is similar to other comparable studies. The 45 and Up Study cohort may, therefore, represent a slightly healthier cohort than the general population, with implications for the generalizability of findings that stem from it. Unfortunately, most cohort studies are vulnerable to this healthy volunteer effect. However, we believe that the use of routinely collected data to follow up participants is an excellent alternative to active longitudinal follow-up for minimizing this potential bias over time. Moreover, internal comparisons within the 45 and Up Study cohort have been shown to be generalizable even in the presence of any selection bias [30]. Finally, although not necessarily a limitation, it is important to remember that the different data sources included in the EXTEND45 Study dataset provide varying levels of population coverage; although government datasets typically capture health service use in both the private and public health care sector, the presence of multiple laboratory service providers in NSW means that laboratory data for the participants are not comprehensive over time or across individuals.

Conclusions

The EXTEND45 Study is, to our knowledge, the first linked data study in Australia to incorporate data from community laboratory service providers with other administrative datasets for chronic disease research. The study represents a unique collaboration among academics, clinicians, primary care researchers, data linkage experts, and laboratory service providers to construct an invaluable data resource that can be used to answer a range of questions relating to chronic disease. Although CKD, diabetes, and CVD have been identified as initial conditions to be examined, consideration of other chronic diseases such as cancer, dyslipidemia, and musculoskeletal conditions is also possible.

Acknowledgments

The EXTEND45 Study is in part funded through peer-reviewed, unconditional grants, including an NSW Cardiovascular Research Network Development Project Grant from the National Heart Foundation of Australia (award no. 100720), a seeding GENESIS grant supported by Roche Products Pty Ltd. (grant no. FR-MIR-0095), and a Rebecca L Cooper Grant (grant no. REB002). In addition, the study was supported in part by unconditional research grants from Eli Lilly (Australia) Pty Ltd. (grant no. PO4100294024), Merck Sharp & Dohme (Australia) Pty Ltd. (grant no. MSD_EXTEND45), and Amgen (Australia) Pty Ltd. (grant no. Amgen01). None of the funding bodies were involved in the design of the study and analysis and interpretation of the study. The study sponsors were not involved in the content development of the EXTEND45 Study. The opinions expressed in this paper are those of the authors and do not necessarily represent those of Eli Lilly (Australia) Pty Ltd, Merck Sharp & Dohme (Australia) Pty Ltd, or Amgen (Australia) Pty Ltd.

This research was completed using data collected through the 45 and Up Study. The 45 and Up Study is managed by the Sax Institute in collaboration with major partner Cancer Council NSW, and partners the National Heart Foundation of Australia (NSW Division), NSW Ministry of Health, NSW Government Family & Community Services—Ageing, Carers, and the Disability Council NSW, and the Australian Red Cross Blood Service. The authors thank the many thousands of people participating in the 45 and Up Study. In addition, the authors thank Services Australia for their supply of MBS and PBS data; NSW Ministry of Health for providing data held in the NSW CCR, NSW APDC, and RBDM; and the ANZDATA Registry and laboratory service providers who have contributed their data to this study. The NSW Center for Record Linkage and the Sax Institute performed the data linkage. The datasets generated during this study are not publicly available because of ethical restrictions. However, upon reasonable request, and with permission of the relevant data custodians, data may be accessed through the SURE. The appropriate steps for becoming a SURE user and accessing SURE will need to be undertaken. Readers can visit the relevant page on the 45 and Up Study website for more information [26].

Authors' Contributions

CF and M Jardine conceived the idea of the study. CF wrote the first draft of the manuscript. CF, CH, LS, M Jun, and M Jardine contributed substantially to the revision of the manuscript. KR led the construction of the database. M Jardine led the team that

created the database. All authors contributed to the design of the database, critically reviewed the manuscript, and have read and approved the final manuscript.

Conflicts of Interest

CC is the Asia Pacific Scientific Leader for Vertis, has received sponsorship from Merck Sharp & Dohme and has been a speaker for Bayer. MG has received honoraria from Shire Pharmaceuticals and Amgen Pty Ltd. SJ is a member of the Board of Directors at The Sax Institute. CP is the current Chair of: Kidney Health Australia, NSW Bureau Health Information and the NSW Cardiovascular Research Network; is a member of the International Advisory Board for AstraZeneca; is a member of Local Advisory Boards for Vifor, Merck Sharpe & Dohm, Boehringer Ingelheim and Otsuka; serves on the Scientific Advisory Board of Pharmaxis; has received travel and accommodation support from Amgen, AstraZeneca and Roche; and has received speaker support from Amgen, AstraZeneca, Novartis and Vifor. DS has received research support from Amgen, Sanofi, Pfizer, Regeneron, Amrin, AstraZeneca and Novartis and provided consultancy services to Amgen, Arrowhead Pharmaceuticals and Sanofi. SZ has participated in the Advisory Board, expert committees, or educational meetings for Boehringer Ingelheim, Eli Lilly, Sanofi, Servier, AstraZeneca, Novo Nordisk, and Merck Sharp & Dohme on behalf of Monash University, with no direct financial compensation. MJu has received unrestricted grant support from VentureWise (a wholly own commercial subsidiary of NPS MedicineWise) to conduct a commissioned project funded by AstraZeneca. MJa is supported by a Medical Research Future Fund Next Generation Clinical Researchers Program Career Development Fellowship; is responsible for research projects that have received unrestricted funding from Gambro, Baxter, CSL Behring, Amgen, Eli Lilly, and Merck; has served on advisory boards sponsored by Akebia, Baxter, Boehringer Ingelheim and Vifor, spoken at scientific meetings sponsored by Janssen, Amgen and Roche; with any consultancy, honoraria or travel support paid to her institution. CF, CH, LS, ACam, AK, TY, ACas, EC, JK, BL, MM, DP, GW, and KR have no competing interests to declare.

Multimedia Appendix 1
Steering committee members of the EXTEND45 Study.
[DOCX File , 17 KB - resprot_v9i4e15646_app1.docx]

References

1. Reddy DV, Gunasekar A. Chronic kidney disease in two coastal districts of Andhra Pradesh, India: role of drinking water. Environ Geochem Health 2013 Aug;35(4):439-454. [doi: 10.1007/s10653-012-9506-7] [Medline: 23475496]
2. Alonso J, Ferrer M, Gandek B, Ware JE, Aaronson NK, Mosconi P, IQOLA Project Group. Health-related quality of life associated with chronic conditions in eight countries: results from the International Quality of Life Assessment (IQOLA) Project. Qual Life Res 2004 Mar;13(2):283-298. [doi: 10.1023/b:qure.0000018472.46236.05] [Medline: 15085901]
3. Fox KA, Steg PG, Eagle KA, Goodman SG, Anderson FA, Granger CB, GRACE Investigators. Decline in rates of death and heart failure in acute coronary syndromes, 1999-2006. J Am Med Assoc 2007 May 2;297(17):1892-1900. [doi: 10.1001/jama.297.17.1892] [Medline: 17473299]
4. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu C. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med 2004 Sep 23;351(13):1296-1305. [doi: 10.1056/NEJMoa041031] [Medline: 15385656]
5. Young A, Lowe J, Byles J, Patterson A. Trends in health service use for women in Australia with diabetes. Aust N Z J Public Health 2005 Oct;29(5):422-428. [doi: 10.1111/j.1467-842x.2005.tb00221.x] [Medline: 16255443]
6. Australian Bureau of Statistics. 2013. Australian Health Survey: Biomedical Results for Chronic Diseases, 2011-12 URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0052011-12 [accessed 2019-04-25]
7. Australian Bureau of Statistics. 2018. National Health Survey: First Results, 2017-18 URL: https://www.abs.gov.au/ausstats/abs@.nsf/mf/4364.0.55.001 [accessed 2019-04-25]
8. Baker Institute. 2012. Diabetes: The Silent Pandemic and Its Impact on Australia URL: https://baker.edu.au/impact/advocacy/the-silent-pandemic [accessed 2019-04-25]
9. Shlipak MG, Katz R, Sarnak MJ, Fried LF, Newman AB, Stehman-Breen C, et al. Cystatin C and prognosis for cardiovascular and kidney outcomes in elderly persons without chronic kidney disease. Ann Intern Med 2006 Aug 15;145(4):237-246. [doi: 10.7326/0003-4819-145-4-200608150-00003] [Medline: 16908914]
10. Australian Bureau of Statistics. 2015. National Health Survey: First Results, 2014-15 URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012014-15?OpenDocument [accessed 2019-04-25]
11. Stevens LA, Li S, Wang C, Huang C, Becker BN, Bomback AS, et al. Prevalence of CKD and comorbid illness in elderly patients in the United States: results from the Kidney Early Evaluation Program (KEEP). Am J Kidney Dis 2010 Mar;55(3 Suppl 2):S23-S33 [FREE Full text] [doi: 10.1053/j.ajkd.2009.09.035] [Medline: 20172445]
12. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015 Aug 22;386(9995):743-800 [FREE Full text] [doi: 10.1016/S0140-6736(15)60692-4] [Medline: 26063472]

13. Shen Y, Cai R, Sun J, Dong X, Huang R, Tian S, et al. Diabetes mellitus as a risk factor for incident chronic kidney disease and end-stage renal disease in women compared with men: a systematic review and meta-analysis. Endocrine 2017 Jan;55(1):66-76. [doi: 10.1007/s12020-016-1014-6] [Medline: 27477292]
14. Gajjala PR, Sanati M, Jankowski J. Cellular and molecular mechanisms of chronic kidney disease with diabetes mellitus and cardiovascular diseases as its comorbidities. Front Immunol 2015;6:340 [FREE Full text] [doi: 10.3389/fimmu.2015.00340] [Medline: 26217336]
15. Chang YT, Wu JL, Hsu CC, Wang JD, Sung JM. Diabetes and end-stage renal disease synergistically contribute to increased incidence of cardiovascular events: a nationwide follow-up study during 1998-2009. Diabetes Care 2014;37(1):277-285. [doi: 10.2337/dc13-0781] [Medline: 23920086]
16. Rivellese AA, Riccardi G, Vaccaro O. Cardiovascular risk in women with diabetes. Nutr Metab Cardiovasc Dis 2010 Jul;20(6):474-480. [doi: 10.1016/j.numecd.2010.01.008] [Medline: 20621459]
17. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, Jafar TH, Heerspink HJ, Mann JF, et al. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. Lancet 2013 Jul 27;382(9889):339-352. [doi: 10.1016/S0140-6736(13)60595-4] [Medline: 23727170]
18. Hur SH, Won KB, Kim IC, Bae JH, Choi DJ, Ahn YK, DIAMOND investigators. Comparison of 2-year clinical outcomes between diabetic versus nondiabetic patients with acute myocardial infarction after 1-month stabilization: Analysis of the prospective registry of DIAMOND (DIabetic acute myocardial infarctiON Disease) in Korea: an observational registry study. Medicine (Baltimore) 2016 Jun;95(25):e3882 [FREE Full text] [doi: 10.1097/MD.0000000000003882] [Medline: 27336875]
19. Navarro MA, Gosch KL, Spertus JA, Rumsfeld JS, Ho PM. Chronic kidney disease and health status outcomes following acute myocardial infarction. J Am Heart Assoc 2016 May 23;5(5):e002772 [FREE Full text] [doi: 10.1161/JAHA.115.002772] [Medline: 27217497]
20. Tonelli M, Muntner P, Lloyd A, Manns BJ, Klarenbach S, Pannu N, Alberta Kidney Disease Network. Risk of coronary events in people with chronic kidney disease compared with those with diabetes: a population-level cohort study. Lancet 2012 Sep 1;380(9844):807-814. [doi: 10.1016/S0140-6736(12)60572-8] [Medline: 22717317]
21. 45Up Study Collaborators, Banks E, Redman S, Jorm L, Armstrong B, Bauman A, et al. Cohort profile: the 45 and up study. Int J Epidemiol 2008 Oct;37(5):941-947 [FREE Full text] [doi: 10.1093/ije/dym184] [Medline: 17881411]
22. The Sax Institute. Questionnaires URL: https://www.saxinstitute.org.au/our-work/45-up-study/questionnaires/ [accessed 2020-01-27]
23. Centre for Health Record Linkage. Validation Studies URL: http://www.cherel.org.au/validation-studies [accessed 2019-04-25]
24. Centre for Health Record Linkage. 2009 Aug. Centre for Health Record Linkage - The First Three Years: 2006-07 to 2008-09 URL: http://www.cherel.org.au/media/13571/cherel-the-first-three-years-2006-07-to-2008-09.pdf [accessed 2019-04-25]
25. Levey AS, Stevens LA, Schmid CH, Zhang Y, Castro AF, Feldman HI, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009 May 5;150(9):604-612 [FREE Full text] [doi: 10.7326/0003-4819-150-9-200905050-00006] [Medline: 19414839]
26. The Sax Institute. Using SURE URL: https://www.saxinstitute.org.au/our-work/sure/using-sure/ [accessed 2019-04-25]
27. Saif MW, Makrilia N, Zalonis A, Merikas M, Syrigos K. Gastric cancer in the elderly: an overview. Eur J Surg Oncol 2010 Aug;36(8):709-717. [doi: 10.1016/j.ejso.2010.05.023] [Medline: 20542657]
28. Hemmelgarn BR, Clement F, Manns BJ, Klarenbach S, James MT, Ravani P, et al. Overview of the Alberta Kidney Disease Network. BMC Nephrol 2009 Oct 19;10:30 [FREE Full text] [doi: 10.1186/1471-2369-10-30] [Medline: 19840369]
29. Faruque LI, Hemmelgarn BR, Wiebe N, Manns BJ, Ravani P, Klarenbach S, Alberta Kidney Disease Network. Factors associated with initiation of chronic renal replacement therapy for patients with kidney failure. Clin J Am Soc Nephrol 2013 Aug;8(8):1327-1335 [FREE Full text] [doi: 10.2215/CJN.10721012] [Medline: 23833317]
30. Mealing NM, Banks E, Jorm LR, Steel DG, Clements MS, Rogers KD. Investigation of relative risk estimates from studies of the same population with contrasting response rates and designs. BMC Med Res Methodol 2010 Apr 1;10:26 [FREE Full text] [doi: 10.1186/1471-2288-10-26] [Medline: 20356408]

Abbreviations

AKDN: Alberta Kidney Disease Network
ANZDATA: Australian and New Zealand Dialysis and Transplant
ATC: Anatomical Therapeutic Chemical
CCR: Central Cancer Registry
CHeReL: Center for Health Record Linkage
CKD: chronic kidney disease
CPRD: Clinical Practice Research Datalink
CVD: cardiovascular disease

DHS: Department of Human Services
EXTEND45: EXamining ouTcomEs in chroNic Disease in the 45 and Up Study
HbA_{1c}: glycated hemoglobin
HREC: Human Research Ethics Committee
ICD-10-AM: International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification
MBS: Medicare Benefits Schedule
NSW: New South Wales
NSW APDC: New South Wales Admitted Patient Data Collection
PBS: Pharmaceutical Benefits Scheme
PHSREC: Population and Health Services Research Ethics Committee
PPN: project-specific person number
RBDM: Registry of Births, Deaths, and Marriages
RRT: renal replacement therapy
SURE: Secure Unified Research Environment

Edited by G Eysenbach; submitted 28.07.19; peer-reviewed by M Harris, S Li, K Jones; comments to author 26.09.19; revised version received 06.12.19; accepted 07.01.20; published 14.04.20.

Please cite as:
Foote C, Hockham C, Sukkar L, Campain A, Kang A, Young T, Cass A, Chow CK, Comino E, Gallagher M, Jan S, Knight J, Liu B, McNamara M, Peiris D, Pollock C, Sullivan D, Wong G, Zoungas S, Rogers K, Jun M, Jardine M
EXamining ouTcomEs in chroNic Disease in the 45 and Up Study (the EXTEND45 Study): Protocol for an Australian Linked Cohort Study
JMIR Res Protoc 2020;9(4):e15646
URL: https://www.researchprotocols.org/2020/4/e15646
doi:10.2196/15646
PMID:32285803

Protocol

Clinical Application of Radioembolization in Hepatic Malignancies: Protocol for a Prospective Multicenter Observational Study

Thomas Helmberger[1*]; Dirk Arnold[2*]; José I Bilbao[3*]; Niels de Jong[4*]; Geert Maleux[5*]; Anders Nordlund[6*]; Bora Peynircioglu[7*]; Bruno Sangro[8,9,10*]; Ricky A Sharma[11*]; Agnes Walk[4*]

[1]Department of Radiology, Neuroradiology and Minimal-Invasive Therapy, Klinikum Bogenhausen, München, Germany

[2]Oncology and Hematology, Asklepios Tumorzentrum Hamburg, Asklepios Klinik Altona, Hamburg, Germany

[3]Interventional Radiology, Clinica Universidad de Navarra, Pamplona, Spain

[4]Cardiovascular and Interventional Radiological Society of Europe, Vienna, Austria

[5]Radiologie, Universitair Ziekenhuis Leuven, Leuven, Belgium

[6]Trial Form Support Aktiebolag, Lund, Sweden

[7]Department of Radiology, School of Medicine, Hacettepe University, Ankara, Turkey

[8]Liver Unit, Clínica Universidad de Navarra, Pamplona, Spain

[9]Instituto de Investigación Sanitaria de Navarra, Pamplona, Spain

[10]Centro de Investigación Biomédica en Red Enfermedades Hepáticas y Digestivas, Pamplona, Spain

[11]National Institute for Health Research University College London Hospitals Biomedical Research Centre, University College London Cancer Institute, University College London, London, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Niels de Jong
Cardiovascular and Interventional Radiological Society of Europe
Neutorgasse 9
Vienna, 1010
Austria
Phone: 43 1904200347
Email: dejong@cirse.org

Abstract

Background: Radioembolization, also known as transarterial radioembolization or selective internal radiation therapy with yttrium-90 (90Y) resin microspheres, is an established treatment modality for patients with primary and secondary liver tumors. However, large-scale prospective observational data on the application of this treatment in a real-life clinical setting is lacking.

Objective: The main objective is to collect data on the clinical application of radioembolization with 90Y resin microspheres to improve the understanding of the impact of this treatment modality in its routine practice setting.

Methods: Eligible patients are 18 years or older and receiving radioembolization for primary and secondary liver tumors as part of routine practice, as well as have signed informed consent. Data is collected at baseline, directly after treatment, and at every 3-month follow-up until 24 months or study exit. The primary objective of the Cardiovascular and Interventional Radiological Society of Europe Registry for SIR-Spheres Therapy (CIRT) is to observe the clinical application of radioembolization. Secondary objectives include safety, effectiveness in terms of overall survival, progression-free survival (PFS), liver-specific PFS, imaging response, and change in quality of life.

Results: Between January 2015 and December 2017, 1047 patients were included in the study. The 24-month follow-up period ended in December 2019. The first results are expected in the third quarter of 2020.

Conclusions: The CIRT is the largest observational study on radioembolization to date and will provide valuable insights to the clinical application of this treatment modality and its real-life outcomes.

Trial Registration: ClinicalTrials.gov NCT02305459; https://clinicaltrials.gov/ct2/show/NCT02305459

International Registered Report Identifier (IRRID): DERR1-10.2196/16296

(JMIR Res Protoc 2020;9(4):e16296) doi:10.2196/16296

KEYWORDS

hepatocellular carcinoma; metastasis; observational study; registries; therapeutic embolization; liver; yttrium-90; radioisotope brachytherapy

Introduction

Primary hepatic malignancies are among the most common cancers of solid organs worldwide and are the fourth to fifth leading cause of death. From the primary liver diseases cirrhosis and hepatocellular carcinoma (HCC), about 180,000 and 75,000 patients, respectively, will die per year in Europe [1,2]. Significant rises in incidence rates for hepatic inflammation, fibrosis, and cirrhosis—predisposing factors for HCC—are expected for the next decade with causes such as autoimmune disease, drug-related effects, or nonalcoholic fatty liver disease [2]. About 4 million people per year are affected by cancer in Europe; 30% to 80% of these patients might develop hepatic metastases [3].

Curative treatment of a hepatic malignancy by liver transplantation, resection, and local ablation can be applied in only about 10% to 25% of cases. Unfortunately, the vast majority of patients do not qualify for these therapies. Various kinds of systemic treatments including chemotherapy, biological therapy, and cancer immunotherapy are offered to this group of patients [4-7]. There is a substantial subset of patients with liver-limited disease that are not suitable for surgical or percutaneous ablative therapies, who experienced early recurrences or no response, significant side effects, or intolerance when treated with systemic therapies. In this setting, transarterial therapies such as chemoembolization, chemoperfusion, or radioembolization (eg, selective internal radiation therapy or transarterial radioembolization) may offer substantial therapeutic improvement [4,7]. Furthermore, these therapies may allow a reduction of systemic side effects, extend periods of freedom from chemotherapy, and prime the liver for other potential local treatment options such as surgery or local ablation [8-11].

Several current, large-scale randomized controlled trials (RCTs) could define the role of radioembolization in first-line and second-line therapy regimens in more advanced primary and secondary liver malignancies [12-15]. Based on this data, current guidelines suggest radioembolization should be considered as a component of the “tool-box” [5,6] in the treatment of HCC, cholangiocarcinoma, and metastatic colorectal cancer (mCRC) following progression after standard therapies or intolerance to other therapies [4,16].

Nevertheless, there is still limited information and understanding of the application of radioembolization in a real-life clinical setting [13,17]. Recent observational studies conducted in the United Kingdom describe the outcome of the real-life application of radioembolization in patients with colorectal liver metastases and intrahepatic cholangiocarcinoma [18,19]. However, similar real-world data from other countries and from patients with other liver malignancies such as HCC are needed by physicians, patients, and paying bodies. More observational data on the use of radioembolization in various clinical settings would elucidate the position of radioembolization in clinical practice and additionally provide data for less established uses of radioembolization, as in metastatic liver disease from other primaries (eg, breast cancer, malignant melanoma, or pancreatic cancer) [20,21].

To further improve the understanding of the real-life clinical application of radioembolization and its impact on clinical practice, the Cardiovascular and Interventional Radiological Society of Europe (CIRSE) initiated the *CIRSE Registry for SIR-Spheres Therapy (CIRT)* for patients treated with radioembolization with yttrium-90 (90Y) resin microspheres (SIR-Spheres, Sirtex Medical Pty Limited; St. Leonards, NSW, Australia). CIRT will collect data on how radioembolization is embedded in real-life clinical practice as well as effectiveness, safety, technical considerations, and patient-reported quality of life (QOL).

Methods

Study Design and Objectives

CIRT is a prospective, multicenter, single-device, noninterventional study of patients with liver tumors treated with radioembolization in daily routine practice as a standard of care. The primary objective is to observe the real-life clinical application of radioembolization and the impact of the treatment in clinical practice. This objective is described by: type of liver cancer, intention of treatment, prior hepatic procedures, associated systemic therapy, and postradioembolization hepatic procedures.

Secondary objectives are effectiveness, safety, technical considerations, and patient-reported outcome measures (see Textbox 1).

Textbox 1. Secondary end points.

Effectiveness end points

- Overall survival
- Progression-free survival
- Hepatic progression-free survival (ie, liver-specific progression-free survival)
- Imaging response

Safety end points

- Day-of-treatment complications
- Adverse events
- Laboratory values

Technical considerations end points

- Patient-related characteristics
- Treatment-related characteristics
- Treatment administration
- Procedure-related outcomes

Patient-reported outcome end points

- Quality of life questionnaire C30
- Additional hepatocellular carcinoma module for patients with hepatocellular carcinoma

Site Selection and Patient Enrollment

Sites that incorporated radioembolization in their standard of care armamentarium and had a minimum amount of experience with the procedures were considered for invitations to participate in the study (ie, 10 or more treatments in the last 12 months and a career history of at least 40 cases). The selected sites were limited to centers in the European Union, Switzerland, Turkey, and Israel. A multidisciplinary steering committee containing experts from the field of medical oncology, diagnostic and interventional radiology, hepatology, surgery, and nuclear medicine developed a list of sites that met the selection criteria and continued to update this list throughout the course of the study. All sites that met the selection criteria were invited directly by the steering committee.

Eligible patients consist of any adult patients treated with radioembolization with 90Y resin microspheres for primary or secondary liver tumors that have signed the informed consent form. No formal sample size calculation was made. The steering committee reasoned that about 1000 patients could be recruited during a period of 3 years, and this would be sufficient to observe the real-life clinical application of radioembolization.

Site enrollment was from August 2014 until April 2017. Patients were included from January 2015 until December 2017. Follow-up data collection ended in December 2019.

This study was performed in accordance with the Declaration of Helsinki and Good Clinical Practice Guidelines. This study was approved by the local ethics committees of participating centers.

Data Collection

Patients enter the study by accepting participation and signing the informed consent form. Baseline data is collected upon allocation. Data on how the treatment was performed and its day-of-treatment outcomes are collected on the day of treatment. The patient is then followed for a maximum of 24 months or until study exit (see Figure 1). Guidelines for radioembolization advise that posttreatment assessments should be performed every 3 months [22]. However, the final decisions on treatment and follow-up schedules are determined by the site-specific medical teams.

Baseline assessments included patient demographics, medical history related to the underlying disease and cancer treatment, as well as cancer type and stage for patients with HCC (Table 1). The technical aspects that may determine how radioembolization is performed or impacts procedure-related outcomes were documented at the time of treatment. Time to event end points and safety end points were collected at every follow-up. In addition to demographic and clinical data, patient-reported outcomes are measured using the European Organisation for the Research and Treatment of Cancer (EORTC) quality of life questionnaire C30 (QLQ-C30) [23,24]. The questionnaire was provided before the treatment, shortly after the treatment, and at every follow-up. When this was not feasible on site, patients were approached via letters or phone calls organized by the study center responsible for the patient. For patients with HCC, in addition to the QLQ-C30, the HCC module was provided to assess factors related to chronic liver disease, as well as issues related to the primary tumor and its treatment [25]. The QLQ-C30 and HCC module were made

available to each patient in their local language, using questionnaires translated and validated by EORTC [26-28].

Demographic and clinical data obtained from medical examinations, medical records, and QOL forms were transferred into an electronic case report form (e-CRF). Patient data was pseudonymized by each site without a centralized pseudonymization policy, and data collection was done according to the General Data Protection Regulation and stored in an encrypted form on a state-of-the-art server in Vienna, Austria. The e-CRF was accessible through a custom-built online database developed by ConexSys Inc (Lincoln, RI) and installed on the CIRSE server in Vienna, Austria. Access to the system is password protected. Statistical analysis will be performed by an independent statistician.

Data quality was ensured through regular remote monitoring of data entered into the database. Remote monitoring includes regular information on data quality (query reports), inclusion of patients, and follow-ups. No on-site monitoring or source data verification was performed due to limited resources.

Figure 1. Flowchart of patient inclusion and measurement time points. CIRT: Cardiovascular and Interventional Radiological Society of Europe Registry for SIR-Spheres Therapy.

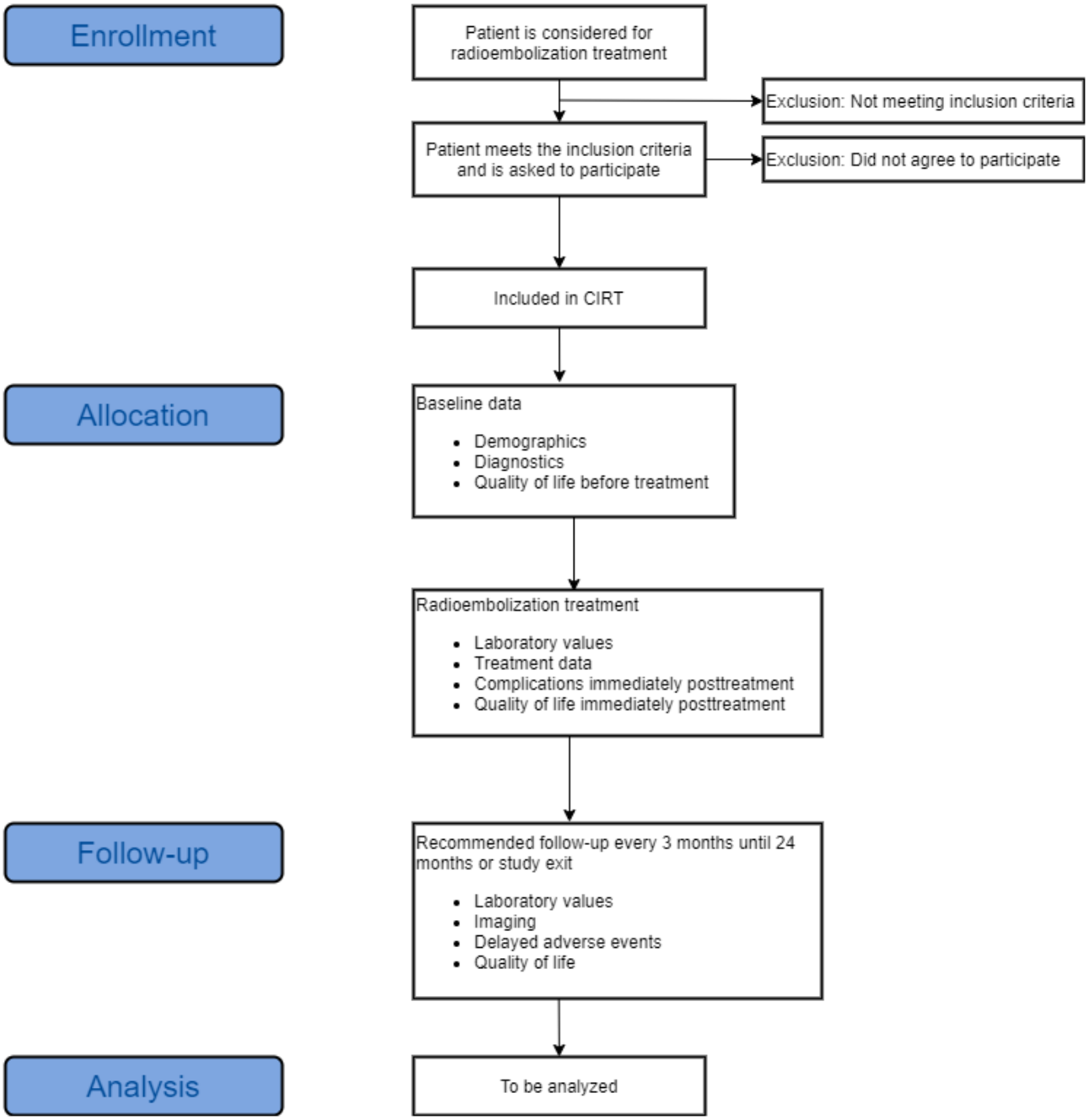


Table 1. Time of measurement for each end point and associated measurements.

End point	Baseline measurement	Day of treatment measurement	Follow-up measurement
Real-life application of radioembolization	• Type of liver cancer • Prior hepatic procedures • Associated systemic therapy (prior systemic therapy)	• Intention of treatment	• Postradioembolization hepatic procedures • Postradioembolization systemic chemotherapy
Effectiveness end points	N/A[a]	• Treatment date	• Date of effectiveness event (OS[b], PFS[c], hepatic PFS, IR[d])
Safety end points	N/A	• Severe day of treatment complications (Grade 3-4)[e]	• Adverse events (Grade 1-5)[e] • Abnormal laboratory values
Technical considerations end points	• Patient-related characteristics • Prior hepatic procedures • Physical characteristics (body surface area, lung shunt study [%], Eastern Cooperative Oncology Group performance status)	• Treatment planning • Treatment administration • Procedure-related outcomes	N/A
Patient-reported outcome end points	• EORTC[f] QLQ-C30[g] • Additional module for HCC[h]	N/A	• EORTC QLQ-C30 • Additional module for HCC

[a]N/A: not applicable.

[b]OS: overall survival.

[c]PFS: progression-free survival.

[d]IR: imaging response.

[e]Grading according to Common Terminology Criteria for Adverse Events version 4.03.

[f]EORTC: European Organisation for the Research and Treatment of Cancer.

[g]QLQ-C30: Quality of Life Questionnaire C30.

[h]HCC: hepatocellular carcinoma.

Statistical Analysis

All patients who were found eligible, enrolled in the registry, underwent radioembolization treatment, and have the minimum amount of data required will be included in analysis. Patients for which the amount of data collected is too scarce to warrant meaningful analysis will be excluded. The number of excluded patients and reason for exclusion will be reported.

Data regarding the primary end point, safety, and technical considerations will be presented by summaries (eg, counts, means, standard deviations) and descriptive statistics. The number of missing observations will be given in all summary tables. Time to event end points will be described graphically through the Kaplan-Meier analysis (including 95% CIs for median survival). The Cox regression will be used to assess the impact of the covariates for time to event end points (Table 2).

Hazard ratios and their 95% CIs will be presented together with *P* values. Sensitivity analyses of time to event data will be performed where censored patients with less than 6 weeks of follow-up will be included as worst case (ie, death or progression).

Table 2. Covariates for time to event end points.

Covariate	Variables
Age (years)	≤69 and ≥70
Sex	Male, female
Number of lines of previous chemotherapy	0, 1, 2-5, ≥6
Primary tumor in situ	Yes, no
Eastern Cooperative Oncology Group performance status	0, 1, and ≥2
Presence of extrahepatic metastases	Yes, no
Prior liver procedures	Yes, no
Number of liver tumors	1, 2-5, and ≥6
Percentage tumor to liver volume	Continuous

QOL is measured with the EORTC QLQ-C30, a questionnaire developed to assess the QOL in patients with cancer. The QLQ-C30 includes five functional scales, three symptom scales, a global health status scale, and six single items. The functional and symptom status of the QLQ-C30 will be assessed according to the scoring manual developed and validated by EORTC [29]. Patients who die during follow-up will be included in the QOL analyses until their date of death. The focus of these analyses will be on patients who are alive at the time point of analysis.

To document the safety and effectiveness of radioembolization and subsequent surgical or interventional procedures, the data will be broken down into subgroups as defined for the covariate analyses above.

Missing data can be expected in the event of a study site failing to enter certain data into the e-CRF, or a patient's withdrawal of consent or being lost to follow-up. Given the descriptive and exploratory nature of CIRT, all available data will be used, and no imputations of missing data will be made. The amount of missing data will be summarized and the subset of patients with missing data will be compared with patients included in the analysis.

In the effectiveness analyses, patients withdrawing consent or being lost to follow-up will be censored at the last time they were observed as being alive (overall survival) or at the time of their last magnetic resonance imaging scan from which disease progression could have been determined (progression-free survival [PFS]). Patients who die during the course of their follow-up in the registry will be regarded as having disease progression in the analyses of PFS and liver-specific PFS.

Missing data for QOL will be handled in line with the suggestions in the EORTC scoring manual for QLQ-C30; the amount of and reasons for completely missing forms will be analyzed and multiple imputations will be used for a part of the completely missing forms (ie, forms considered to be missing at random or due to administrative failure) [29].

Results

Between August 2014, and April 2017, 63 hospitals that met the selection criteria were invited, 36 (57%) of which were activated (ie, were trained, had access to database, were eligible to include patients), of which 29 (81%) enrolled patients.

A total of 1047 patients were collected during the trial period between January 1, 2015, and December 31, 2017. Follow-up data collection ended on December 31, 2019. Final results are expected to be published in 2020 or 2021.

Discussion

Summary

CIRT is the first European-wide prospective observational study of patients with liver tumors treated with radioembolization as the standard of care and will provide an important addition to the body of knowledge on radioembolization treatment.

Data from several large RCTs on mCRC and HCC were published while CIRT was already ongoing and more can be expected in the future [12,30,31]. However, these studies provide insufficient data with respect to "real life" clinical practice since there may be patient cohorts that benefit more from radioembolization than the patient cohorts included in these trials [32-34]. CIRT was designed to also explore patient cohorts for which less data is available. The writing group hopes that the outcomes of CIRT would provide valuable insights in these less explored patient populations to gain further insight into real world clinical practice of radioembolization. The extent to which this data relates to the overall body of data available on radioembolization will be discussed in the results publication.

Limitations

The single-arm observational study nature of the study design implies that several limitations need to be considered. The absence of a contemporaneous comparator group may limit our interpretation of the clinical data reported. Despite this, important positive features of the study are the long-term data collection for patients included and the valuable information collected on health-related QOL. The size of the study should enhance patient selection by providing new information on the patient subgroups that benefit most from this treatment in day-to-day clinical practice.

Furthermore, selection bias could occur at several stages of the study. An element of site selection was introduced when the Steering Committee created a list of sites that met the selection criteria of minimal expertise with radioembolization treatment. The potential of patient selection bias is addressed by a contractual agreed upon prospective design, whereby sites agreed to present the possibility to participate in the CIRT to all eligible patients consecutively.

Another source for bias relates to drop-out during follow-up. The sites included in the study have been advised to continue patient follow-up at least 24 months after the first treatment. However, patients may be lost to follow-up or withdraw their consent at any time. Loss to follow-up may relate to the course of disease or site procedures; for example, when a patient is treated at a participating site but followed-up by their referring physician, sometimes in another country. It will therefore not always be possible for study sites to collect additional follow-up data and the patient will be reported as lost to follow-up. Baseline characteristics will be compared between those lost to follow-up and those who have follow-up data to assess differences with the potential to bias results. Notable differences on a background characteristic will prompt additional analyses where the sensitivity of the results to that background characteristic will be further assessed.

Conclusion

To date, the CIRT is the largest observational study on the use of radioembolization with 90Y resin microspheres. The objective of the study was to observe the real-life clinical application of radioembolization and its impact in clinical practice. Currently, radioembolization has a position in the guidelines for the treatment of HCC, cholangiocarcinoma, and mCRC, and more data is warranted on the possibilities that this treatment has for

these diseases. Furthermore, additional data is needed for the effects of other lines of therapy and treatment of liver metastatic disease from other primary sites. Observing a large study population may reveal treatment aspects that would justify a wider application of this innovative treatment.

Acknowledgments

This observational study was sponsored and executed by a CIRSE, as a medical society, with an unrestricted grant of the manufacturer of the SIR-Spheres 90Y resin microspheres. Thanks to the financial support from SIRTEX Medical Europe GmbH (Bonn, Germany), CIRSE was able to act as a scientifically independent sponsor of the study with sole ownership of the data. The electronic data capturing system was developed and supported by ConexSys Inc (Lincoln, RI). Necessary server infrastructure was developed and maintained by ITEA GmbH (Vienna, Austria). The authors thank the CIRSE Central Office and the staff of the CIRSE Clinical Research Department for their support during the design and the setup of the study and the drafting of the manuscript.

Conflicts of Interest

Author TH received speaker honoraria from SIRTEX Medical Europe. Author DA received consulting fees and speaker honoraria from TERUMO, Boston Scientific, SIRTEX Medical Europe, and Biocompatibles. Author JB received speaker honoraria and consultation fees from SIRTEX Medical Europe. Author BS received scientific grants from Bristol-Myers-Squibb and SIRTEX Medical Europe; consultation fees and speaker honoraria from Astra Zeneca, Bayer, Bristol-Myers-Squibb, Ipsen, SIRTEX Medical Europe, and Terumo; consultation fees from Adaptimmune, BTG, Eli Lilly, and Onxeo; and fees for data review activities from Roche and BTG. Author RS received scientific grants from SIRTEX Technology and consultation fees from SIRTEX Medical Ltd. Author GM received speaker fees from SIRTEX Medical Europe. Authors NdJ, AN, BP, and AW reported no conflicts of interest.

References

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019 Apr 15;144(8):1941-1953. [doi: 10.1002/ijc.31937] [Medline: 30350310]
2. Blachier M, Leleu H, Peck-Radosavljevic M, Valla D, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. J Hepatol 2013 Mar;58(3):593-608 [FREE Full text] [doi: 10.1016/j.jhep.2012.12.005] [Medline: 23419824]
3. de Ridder J, de Wilt JH, Simmer F, Overbeek L, Lemmens V, Nagtegaal I. Incidence and origin of histologically confirmed liver metastases: an explorative case-study of 23,154 patients. Oncotarget 2016 Aug 23;7(34):55368-55376 [FREE Full text] [doi: 10.18632/oncotarget.10552] [Medline: 27421135]
4. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. Ann Oncol 2016 Aug;27(8):1386-1422 [FREE Full text] [doi: 10.1093/annonc/mdw235] [Medline: 27380959]
5. Valle JW, Borbath I, Khan SA, Huguet F, Gruenberger T, Arnold D, ESMO Guidelines Committee. Biliary cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2016 Sep;27(suppl 5):v28-v37. [doi: 10.1093/annonc/mdw324] [Medline: 27664259]
6. Giammarile F, Bodei L, Chiesa C, Flux G, Forrer F, Kraeber-Bodere F, Therapy, Oncology and Dosimetry Committees. EANM procedure guideline for the treatment of liver cancer and liver metastases with intra-arterial radioactive compounds. Eur J Nucl Med Mol Imaging 2011 Jul;38(7):1393-1406. [doi: 10.1007/s00259-011-1812-2] [Medline: 21494856]
7. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatocellular carcinoma. J Hepatol 2018 Jul;69(1):182-236. [doi: 10.1016/j.jhep.2018.03.019] [Medline: 29628281]
8. Garlipp B, de Baere T, Damm R, Irmscher R, van Buskirk M, Stübs P, et al. Left-liver hypertrophy after therapeutic right-liver radioembolization is substantial but less than after portal vein embolization. Hepatology 2014 May;59(5):1864-1873. [doi: 10.1002/hep.26947] [Medline: 24259442]
9. Salem R, Lewandowski RJ, Kulik L, Wang E, Riaz A, Ryu RK, et al. Radioembolization results in longer time-to-progression and reduced toxicity compared with chemoembolization in patients with hepatocellular carcinoma. Gastroenterology 2011 Feb;140(2):497-507.e2 [FREE Full text] [doi: 10.1053/j.gastro.2010.10.049] [Medline: 21044630]
10. Salem R, Gilbertsen M, Butt Z, Memon K, Vouche M, Hickey R, et al. Increased quality of life among hepatocellular carcinoma patients treated with radioembolization, compared with chemoembolization. Clin Gastroenterol Hepatol 2013 Oct;11(10):1358-1365.e1. [doi: 10.1016/j.cgh.2013.04.028] [Medline: 23644386]
11. Salem R, Gordon AC, Mouli S, Hickey R, Kallini J, Gabr A, et al. Y90 radioembolization significantly prolongs time to progression compared with chemoembolization in patients with hepatocellular carcinoma. Gastroenterology 2016 Dec;151(6):1155-1163.e2 [FREE Full text] [doi: 10.1053/j.gastro.2016.08.029] [Medline: 27575820]

12. Chauhan N, Mulcahy MF, Salem R, Benson Iii AB, Boucher E, Bukovcan J, et al. TheraSphere yttrium-90 glass microspheres combined with chemotherapy versus chemotherapy alone in second-line treatment of patients with metastatic colorectal carcinoma of the liver: protocol for the EPOCH phase 3 randomized clinical trial. JMIR Res Protoc 2019 Jan 17;8(1):e11545 [FREE Full text] [doi: 10.2196/11545] [Medline: 30664496]
13. Sposito C, Mazzaferro V. The SIRveNIB and SARAH trials, radioembolization vs sorafenib in advanced HCC patients: reasons for a failure, and perspectives for the future. Hepatobiliary Surg Nutr 2018 Dec;7(6):487-489 [FREE Full text] [doi: 10.21037/hbsn.2018.10.06] [Medline: 30652096]
14. Vilgrain V, Pereira H, Assenat E, Guiu B, Ilonca AD, Pageaux G, SARAH Trial Group. Efficacy and safety of selective internal radiotherapy with yttrium-90 resin microspheres compared with sorafenib in locally advanced and inoperable hepatocellular carcinoma (SARAH): an open-label randomised controlled phase 3 trial. Lancet Oncol 2017 Dec;18(12):1624-1636. [doi: 10.1016/S1470-2045(17)30683-6] [Medline: 29107679]
15. Wasan HS, Gibbs P, Sharma NK, Taieb J, Heinemann V, Ricke J, FOXFIRE trial investigators, SIRFLOX trial investigators, FOXFIRE-Global trial investigators, et al. First-line selective internal radiotherapy plus chemotherapy versus chemotherapy alone in patients with liver metastases from colorectal cancer (FOXFIRE, SIRFLOX, and FOXFIRE-Global): a combined analysis of three multicentre, randomised, phase 3 trials. Lancet Oncol 2017 Sep;18(9):1159-1171 [FREE Full text] [doi: 10.1016/S1470-2045(17)30457-6] [Medline: 28781171]
16. European Association For The Study Of The Liver, European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 2012 Apr;56(4):908-943 [FREE Full text] [doi: 10.1016/j.jhep.2011.12.001] [Medline: 22424438]
17. van Hazel GA, Heinemann V, Sharma NK, Findlay MPN, Ricke J, Peeters M, et al. SIRFLOX: randomized phase III trial comparing girst-line mFOLFOX6 (plus or minus bevacizumab) versus mFOLFOX6 (plus or minus bevacizumab) plus selective internal radiation therapy in patients with metastatic colorectal cancer. J Clin Oncol 2016 May 20;34(15):1723-1731. [doi: 10.1200/JCO.2015.66.1181] [Medline: 26903575]
18. White J, Carolan-Rees G, Dale M, Morgan HE, Patrick HE, See TC, et al. Analysis of a national programme for selective internal radiation therapy for colorectal cancer liver metastases. Clin Oncol (R Coll Radiol) 2019 Jan;31(1):58-66. [doi: 10.1016/j.clon.2018.09.002] [Medline: 30297164]
19. White J, Carolan-Rees G, Dale M, Patrick HE, See TC, Bell JK, et al. Yttrium-90 transarterial radioembolization for chemotherapy-refractory intrahepatic cholangiocarcinoma: a prospective, observational study. J Vasc Interv Radiol 2019 Aug;30(8):1185-1192. [doi: 10.1016/j.jvir.2019.03.018] [Medline: 31255499]
20. Kuei A, Saab S, Cho S, Kee ST, Lee EW. Effects of Yttrium-90 selective internal radiation therapy on non-conventional liver tumors. World J Gastroenterol 2015 Jul 21;21(27):8271-8283 [FREE Full text] [doi: 10.3748/wjg.v21.i27.8271] [Medline: 26217079]
21. Barbier CE, Garske-Román U, Sandström M, Nyman R, Granberg D. Selective internal radiation therapy in patients with progressive neuroendocrine liver metastases. Eur J Nucl Med Mol Imaging 2016 Jul;43(8):1425-1431. [doi: 10.1007/s00259-015-3264-6] [Medline: 26631239]
22. Mahnken AH, Spreafico C, Maleux G, Helmberger T, Jakobs TF. Standards of practice in transarterial radioembolization. Cardiovasc Intervent Radiol 2013 Jun;36(3):613-622. [doi: 10.1007/s00270-013-0600-8] [Medline: 23511991]
23. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 03;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
24. Fayers P, Bottomley A, EORTC Quality of Life Group, Quality of Life Unit. Quality of life research within the EORTC-the EORTC QLQ-C30. Eur J Cancer 2002 Mar;38 Suppl 4:S125-S133. [doi: 10.1016/s0959-8049(01)00448-8] [Medline: 11858978]
25. Blazeby JM, Currie E, Zee BCY, Chie W, Poon RT, Garden OJ, EORTC Quality of Life Group. Development of a questionnaire module to supplement the EORTC QLQ-C30 to assess quality of life in patients with hepatocellular carcinoma, the EORTC QLQ-HCC18. Eur J Cancer 2004 Nov;40(16):2439-2444. [doi: 10.1016/j.ejca.2004.06.033] [Medline: 15519517]
26. Scott NW, Etta JA, Aaronson NK, Bottomley A, Fayers PM, Groenvold M, et al. An evaluation of the response category translations of the EORTC QLQ-C30 questionnaire. Qual Life Res 2013 Aug;22(6):1483-1490. [doi: 10.1007/s11136-012-0276-6] [Medline: 23054491]
27. Scott NW, Fayers PM, Bottomley A, Aaronson NK, de Graeff A, Groenvold M, EORTC and the Quality of Life Cross-Cultural Meta-Analysis Group. Comparing translations of the EORTC QLQ-C30 using differential item functioning analyses. Qual Life Res 2006 Aug;15(6):1103-15; discussion 1117. [doi: 10.1007/s11136-006-0040-x] [Medline: 16900290]
28. Cankurtaran ES, Ozalp E, Soygur H, Ozer S, Akbiyik DI, Bottomley A. Understanding the reliability and validity of the EORTC QLQ-C30 in Turkish cancer patients. Eur J Cancer Care (Engl) 2008 Jan;17(1):98-104. [doi: 10.1111/j.1365-2354.2007.00827.x] [Medline: 18181898]
29. Fayers P, Aaronson N, Bjordal K, Groenvold M, Curran D, Bottomley A. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels: European Organisation for Research and Treatment of Cancer; 2001.
30. Chauhan N, Bukovcan J, Boucher E, Cosgrove D, Edeline J, Hamilton B, et al. Intra-Arterial TheraSphere Yttrium-90 Glass Microspheres in the Treatment of Patients With Unresectable Hepatocellular Carcinoma: Protocol for the STOP-HCC

Phase 3 Randomized Controlled Trial. JMIR Res Protoc 2018 Aug 15;7(8):e11234 [FREE Full text] [doi: 10.2196/11234] [Medline: 30111528]
31. Gebski V, Gibbs E, Gandhi M, Chatellier G, Dinut A, Pereira H, et al. VESPRO: an individual patient data prospective meta-analysis of selective internal radiation therapy versus sorafenib for advanced, locally advanced, or tecurrent hepatocellular carcinoma of the SARAH and SIRveNIB Trials. JMIR Res Protoc 2017 Feb 15;6(2):e17 [FREE Full text] [doi: 10.2196/resprot.7016] [Medline: 28202430]
32. Bester L, Meteling B, Pocock N, Pavlakis N, Chua TC, Saxena A, et al. Radioembolization versus standard care of hepatic metastases: comparative retrospective cohort study of survival outcomes and adverse events in salvage patients. J Vasc Interv Radiol 2012 Jan;23(1):96-105. [doi: 10.1016/j.jvir.2011.09.028] [Medline: 22079516]
33. Seidensticker R, Denecke T, Kraus P, Seidensticker M, Mohnike K, Fahlke J, et al. Matched-pair comparison of radioembolization plus best supportive care versus best supportive care alone for chemotherapy refractory liver-dominant colorectal metastases. Cardiovasc Intervent Radiol 2012 Oct;35(5):1066-1073. [doi: 10.1007/s00270-011-0234-7] [Medline: 21800231]
34. Hendlisz A, Van den Eynde M, Peeters M, Maleux G, Lambert B, Vannoote J, et al. Phase III trial comparing protracted intravenous fluorouracil infusion alone or with yttrium-90 resin microspheres radioembolization for liver-limited metastatic colorectal cancer refractory to standard chemotherapy. J Clin Oncol 2010 Aug 10;28(23):3687-3694. [doi: 10.1200/JCO.2010.28.5643] [Medline: 20567019]

Abbreviations

CIRSE: Cardiovascular and Interventional Radiological Society of Europe
CIRT: Cardiovascular and Interventional Radiological Society of Europe Registry for SIR-Spheres Therapy.
e-CRF: electronic case report form
EORTC: European Organization for the Research and Treatment of Cancer
HCC: hepatocellular carcinoma
mCRC: metastatic colorectal cancer
PFS: progression-free survival
QLQ-C30: quality of life questionnaire C30
QOL: quality of life
RCT: randomized controlled trials
SIR: selective internal radiation
90Y: yttrium-90.

Edited by G Eysenbach; submitted 17.09.19; peer-reviewed by J Farzi, D Coffman; comments to author 08.12.19; revised version received 30.01.20; accepted 21.02.20; published 22.04.20.

Please cite as:
Helmberger T, Arnold D, Bilbao JI, de Jong N, Maleux G, Nordlund A, Peynircioglu B, Sangro B, Sharma RA, Walk A
Clinical Application of Radioembolization in Hepatic Malignancies: Protocol for a Prospective Multicenter Observational Study
JMIR Res Protoc 2020;9(4):e16296
URL: http://www.researchprotocols.org/2020/4/e16296/
doi:10.2196/16296
PMID:32319960

Protocol

Monitoring Health and Well-Being in Emerging Adults: Protocol for a Pilot Longitudinal Cohort Study

Reidar P Lystad[1], PhD; Diana Fajardo Pulido[1], MIPH/MHM; Lorna Peters[2], PhD; Melissa Johnstone[3], PhD; Louise A Ellis[1], PhD; Jeffrey Braithwaite[1], PhD; Viviana Wuthrich[2], PhD; Janaki Amin[4], PhD; Cate M Cameron[5], PhD; Rebecca J Mitchell[1], PhD

[1]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

[2]Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney, Australia

[3]Department of Educational Studies, Macquarie University, Sydney, Australia

[4]Department of Health Systems and Populations, Macquarie University, Sydney, Australia

[5]Jamieson Trauma Institute, Metro North Hospital and Health Service, Queensland Health, Brisbane, Australia

Corresponding Author:
Reidar P Lystad, PhD
Australian Institute of Health Innovation
Macquarie University
75 Talavera Road
Sydney, 2109
Australia
Phone: 61 298502464
Email: reidar.lystad@mq.edu.au

Abstract

Background: Emerging adulthood is a unique segment of an individual's life course. The defining features of this transitional period include identity exploration, instability, future possibilities, self-focus, and feeling in-between adolescence and adulthood, all of which are thought to affect quality of life, health, and well-being. A longitudinal cohort study with a comprehensive set of measures would be a unique and valuable resource for improving the understanding of the multi-faceted elements and unique challenges that contribute to the health and well-being of emerging adults.

Objective: The main aim of this pilot study is to evaluate the feasibility and acceptability of recruiting university graduates to establish a longitudinal cohort study to inform our understanding of emerging adulthood.

Methods: This is a pilot longitudinal cohort study of Australian university graduates. It will involve collecting information via online surveys (baseline and 12-month follow-up) and data linkage with health records. Recruitment, response, and retention rates will be calculated. Descriptive analysis of the representativeness of recruited participants and completeness of survey responses will be conducted.

Results: Participant recruitment was completed in October 2018, and data collection for the baseline and follow-up surveys was completed in November 2019. As of April 2020, the process of acquiring health records from administrative data collections has commenced.

Conclusions: The findings from this pilot study will identify areas for improvement and inform the development of a future longitudinal cohort study of emerging adults.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12618001364268; https://tinyurl.com/teec8wh

International Registered Report Identifier (IRRID): DERR1-10.2196/16108

(JMIR Res Protoc 2020;9(4):e16108) doi:10.2196/16108

KEYWORDS

young adults; emerging adulthood; health; well-being; health-related quality of life

Introduction

Throughout young people's lives, there are many events and factors that can affect their life course. Emerging adulthood is the transitional period from late teens through to the late twenties and is characterized by a higher degree of diversity, instability, and uncertainty [1]. The defining features of emerging adulthood include identity exploration (ie, exploring available options for life especially in love and work); instability (ie, being subject to numerous changes and shifting choices); future possibilities (ie, multiple available options where different futures remain possible); self-focus (ie, focus on forming oneself); and feeling in-between (ie, neither an adolescent nor an adult). Demographic norms change considerably during emerging adulthood, especially in terms of residential status and school attendance. In their late teens, most people live with one or more parents and attend school; whereas most people in their thirties work full-time, live independently, and cohabitate with a romantic partner. These features are thought to impact the quality of life and well-being of emerging adults [2].

Transitioning from education to work life can be particularly challenging for emerging adults [3]. While the experience of tertiary education gives the opportunity to explore different identities and lifestyles, work is often more salient in shaping one's identity because of its central role in adult life [2]. Work can be pivotal for long-term prospects, such as acquiring financial independency, career, marriage or partnership, and parenthood [2]. Difficulties in transitioning between education and work life can negatively affect the health and well-being of emerging adults, and unsuccessful transitions can lead to mental health problems later in life [4,5].

The sense of instability, uncertainty, and multitude of future possibilities can negatively impact physical health, health-related quality of life (HRQOL), and well-being of emerging adults [6]. HRQOL is a multi-dimensional concept that purports to quantify the relationships between physical and mental health status and quality of life over time [7]. Many popular HRQOL metrics typically measure self-perceived health status [8]. A related concept is well-being, which evaluates the positive aspects of an individual's life, including life satisfaction and positive emotions [9]. Both HRQOL and well-being have been used to measure the impact of illness and disability on the quality of life of emerging adults. For instance, Pons-Villanueva et al [10] found that university graduates who had been involved in a motor vehicle crash had poorer HRQOL four years post crash. Van Oostrom et al [11] reported that adults who had adopted an active lifestyle experienced better HRQOL over time. Further, Buhl [3] identified that emerging adults who did not go to university reported poorer parent-child relationships compared to those who transitioned from university to work life.

Although several studies have investigated aspects of HRQOL and well-being in emerging adults, more information is needed to better understand these relationships. A more comprehensive view of the multi-faceted elements and unique challenges that contribute to the health and well-being of emerging adults, including education, employment, lifestyle, HRQOL, well-being, social support, life events, carer responsibilities, and use of social media technology is needed. Thus, conducting a large, longitudinal cohort study with a comprehensive set of measures would be a unique and valuable resource for improving our understanding of education, health, and lifestyle factors and their impact on resilience, career trajectories, and lived experiences over a unique segment of an individual's life course.

This pilot cohort study aims to establish the feasibility and acceptability of recruiting university graduates to establish a longitudinal cohort study to inform our understanding of emerging adulthood. Specifically, the study will evaluate: (1) the feasibility of our research methods to recruit university graduates at one large Australian university, including determination of opt-in and opt-out rates for data linkage of health records and survey responses; (2) the representativeness of recruited participants; (3) our ability to obtain baseline survey data, including completion of individual survey instruments; (4) our ability to retain participants and collect follow-up survey data 12 months post baseline, including completion of individual survey instruments; (5) and identify areas for improvement for future studies.

Methods

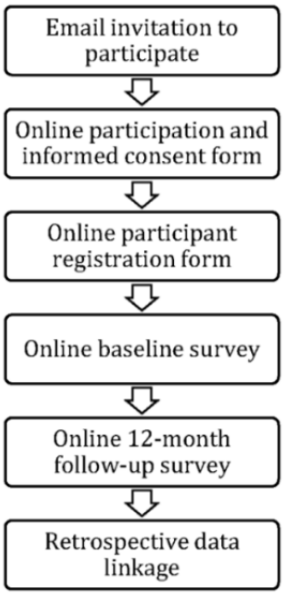
Registration

This study was registered with the Australian New Zealand Clinical Trials Registry (ACTRN) on August 14, 2018 (ACTRN12618001364268).

Design

This is a pilot longitudinal cohort study of Macquarie University graduates. It involves collecting information via online surveys (ie, baseline and 12-month follow-up) and data linkage with health records. A flowchart of the study design is depicted in Figure 1.

Figure 1. Flowchart of the study design.



Recruitment

All students graduating from Macquarie University in 2018 are eligible to participate in the study. Macquarie University is a public university located in Sydney, Australia. The university has five faculties (ie, Faculty of Arts, Faculty of Business and Economics, Faculty of Human Sciences, Faculty of Medicine and Health Sciences, and Faculty of Science and Engineering), which collectively host approximately 45,000 students, including 33,000 undergraduate students, 9000 postgraduate students, and 1500 higher-degree research students. Each year, approximately 7000 students graduate with an undergraduate or postgraduate degree.

Invitations to participate in this study will be distributed to the graduates via email from the Macquarie University Graduation Office in conjunction with the fall and spring graduation ceremonies (ie, April 12-27, 2018 and September 19-28, 2018). The initial email invitation will be sent out during the graduation ceremony period, while three reminder emails will be sent out during the following 4-6 weeks. The email invitations include a brief description of the purpose of the study, what participation will involve, and a link to the *Macquarie University – Monitoring of Injury and Psychosocial Health Outcomes, Career Trajectories and Continuing Education, Lived Experiences, and Social Connectedness* (MQ-MINDS) project website with a full Participant Information and Consent Form. The Participant Information and Consent Form contains details about the purpose of the study, what participation will involve, a description of benefits and risks of taking part in the study, confidentiality and privacy arrangements, funding for the study, and consent to participate.

Participants will be given the opportunity to opt out of having their survey responses linked to their health records (ie, ambulance, emergency department, hospitalization, cancer registry, and mortality records). Participants are then directed to an online participant registration form that securely records their personally identifiable information, including their name, residential address, mobile phone number, and email address. Upon completing the online registration form, participants will receive an email with an individualized link to the baseline survey.

Data Collection

Data will be collected through online surveys and, for those that did not opt out, health record linkage. The online surveys will be administered online via the Qualtrics XM Platform (Qualtrics International Inc) at baseline and at 12 months post baseline. The baseline and 12-month follow-up surveys will comprise the same battery of validated questionnaires and instruments designed to capture data regarding: sociodemographics, education, employment, job satisfaction, mentoring, self-perceived health status, work-life balance, connectedness, resilience, injury, risk behaviors, life events, as well as social media and technology use. It will take approximately 25 minutes to complete the online survey.

Survey Instrument

An overview of the domains and specific questionnaires included in the survey is provided in Table 1. Because the target population is a subset of the general Australian population, the survey is comprised of instruments that are designed or adapted for use in the Australian population, whenever possible. This will facilitate more direct comparison of the data collected in this study with existing normative data from the general Australian population.

Table 1. Overview of survey domains and measures.

Domain	Measures
Sociodemographics	• Standardized questions about gender, sexual orientation, height, weight, ethnicity, language, marital status, house tenure, and income
Tertiary education	• Questions about previous academic qualifications and current enrollment in tertiary education
Employment	• Questions about occupation, employment status, job satisfaction, job barriers, and future employment goals • Questions about career mentoring • Role Balance Scale
Lifestyle	• Questions about physical activity and sedentary behavior
Health	• EuroQoL 5-dimension • Short Form Health Survey • Kessler Psychological Distress Scale • Social Interaction Anxiety Scale • General Anxiety Disorder scale • Questions about smoking, alcohol consumption, drug use, and sexual behavior
Social support	• Multidimensional Scale of Perceived Social Support • Questions about social participation (eg, community, church, or self-help groups)
Life events	• Social Readjustment Rating Scale • Brief Resilience Scale
Carer activities	• Questions about carer responsibilities and activities
Social media and technology	• Questions about access to the internet and devices used (eg, laptop, mobile phone, tablet) • Questions about use of social networking sites (eg, Facebook, Twitter, Snapchat) • Questions about experiences with using social media

Sociodemographics

For the sociodemographic domain, questions about gender, sexual orientation, ethnicity, language, marital status, living arrangement, and household income are derived from the New South Wales Population Health Survey [12].

Tertiary Education

The education domain includes questions about previous academic qualifications, level of previous academic degrees, and current academic programs.

Employment

The employment domain is comprised of three subdomains: general questions about current employment, work-life balance, and career mentoring. In regard to current employment, questions about occupation, employment status, job satisfaction, and future employment goals are adapted from the Australian Workplace Relations study [13]. Job satisfaction is assessed on a 7-point Likert scale ranging from “extremely satisfied” to “extremely dissatisfied,” within seven different perceived aspects of the current job (ie, flexibility, decision making, autonomy, salary, job security, job content, and working conditions) and an overall question: “how satisfied are you with your current job?”. Information on work-life balance is recorded using the Role Balance Scale (RBS) [14]. The RBS consists of eight items that evaluate the engagement of participants across different roles and the ability to incorporate the newly emerging roles within their life. The first five statements focus on the balance and enjoyment across different roles, and the distribution of importance between roles and overall satisfaction. The last three statements indicate the self-perceived importance of each role in the participants' lives. Information about the perceived benefits and potential role of mentors in the participants’ career will be collected using an adapted instrument developed by DeCastro et al [15]. There are nine items that consider different aspects of mentoring relationships (eg, improvement of job performance; mentor perceived as a role model; increased social network; advise to further develop professional career; resources to develop professional career; advise in keeping work-life balance; and develop new knowledge, skills and ethical behavior). These items are assessed on a 7-point Likert scale ranging from “extremely satisfied” to “extremely dissatisfied.”

Lifestyle

Questions about physical activity and sedentary behavior are adapted from the New South Wales Population Health Survey [12]. These include questions about time spent walking and frequency; time spent doing moderate, strengthening, and vigorous activities per week; and time spent sleeping, sitting at work, watching television, and using technology devices such as computers, tablets, or smartphones.

Health

The health domain comprises several instruments assessing various aspects of physical and mental health and HRQOL. The

12-item Short Form Health Survey (SF-12) provides insight regarding the participant's physical and mental health measured through eight dimensions (ie, physical functioning, role physical, role emotional, mental health, body pain, general health, vitality, and social functioning) using a 7-point Likert scale [16]. The SF-12 has demonstrated great feasibility in monitoring the health of specific populations [16].

The EuroQoL 5-dimension (EQ-5D) is a widely used instrument to describe and value health. It comprises two parts: a five-item descriptive system and a visual analogue scale [17]. The five items (ie, mobility, self-care, usual activities, pain/discomfort and anxiety/depression) are rated using five levels: "No problems," "Slight problems," "Moderate problems," "Severe problems," and "Extreme problems." The final question asks respondents to rate their health on a scale ranging from 0 (ie, "best imaginable health state") to 100 (ie, "worst imaginable health state") [18].

The General Anxiety Disorder scale (GAD-7) is a tool with strong validity to identify probable cases of GAD that can be associated with multiple domains of functional impairment and disability days [19]. It consists of seven items that identify symptoms of anxiety (eg, feeling nervous or anxious, being unable to stop or control worrying, worrying too much, having trouble relaxing, being restless, becoming easily annoyed or irritable, and feeling afraid as if something awful might happen) over the past 2 weeks, and rates their severity on a scale from 0 ("Not at all") to 3 ("Nearly every day"). If applicable, the GAD-7 also includes a question about the respondent's perceived difficulty in performing daily activities due to these symptoms [19].

The Social Interaction Anxiety Scale (SIAS-6) is an accurate and efficient psychometric instrument that aims to assess social interaction anxiety as the core feature of social anxiety disorder [20]. The instrument is comprised of six statements about meeting and talking to strangers, friends, or members of the opposite sex. Each statement is rated using a 5-point scale, ranging from 0 ("Not at all characteristic or true of me") to 4 ("Extremely characteristic or true of me"), to reflect the level of general anxiety associated with the initiation and maintenance of social interactions [20].

The Kessler Psychological Distress Scale (K10) is a screening instrument used to determine mental illness in health risk appraisal [21]. The K10 comprises 10 questions about emotional states (eg, feelings of fatigue, motor agitation, guilt, restlessness, anxiety, and depression), each of which is rated on a 5-point scale ranging from 1 ("None of the time") to 5 ("All of the time") [21]. The individual item scores are summed, yielding an overall score ranging from a minimum of 10 to a maximum of 50, with higher scores indicating higher levels of distress. The K10 scores are categorized as "Low" (10-15), "Moderate" (16-21), "High" (22-29), and "Very high" (30-50) [22].

In addition to these instruments, the health domain also includes questions about injury history and health risk behaviors. Questions about the respondent's 12-month history of motor vehicle crash incidents, injury due to external trauma, and injury-related hospitalizations have been adapted from the Seguimiento University of Navarra study [10]. Questions about health risk behaviors, which are derived from the New South Wales Population Health Survey [12], consist of questions about smoking (including use of electronic cigarettes), alcohol consumption, illicit and recreational drug use, and sexual behavior.

Social Support

The social support domain comprises three subdomains: social connectedness, resilience, and perceived social support. The question about social connectedness and community networks of the respondents is adapted from the Nurses' Health Study II [23]. This question explores how often the respondent takes part in social groups such as workgroups, church-connected groups, self-help groups, charity groups, and public service or community groups. The question about resilience is adapted from the Brief Resilience Scale (BRS) [24]. It includes four items that assess the respondent's self-reported ability to look for creative ways to alter difficult situations, control reactions, grow in positive ways by dealing with difficult situations, and ways to replace losses encountered in life [24]. Perceived social support is measured using the Multidimensional Scale of Perceived Social Support (MSPSS) [25]. The MSPSS comprises 12 items that are rated on a 7-point Likert scale ranging from "Strongly disagree" to "Strongly agree." The items are divided into three subscales based on the source of perceived social support (ie, family, friends, and a significant other).

Life Events

The Social Readjustment Rating Scale (SRRS) is used to measure the impact of major life events [26]. The SRRS consists of 43 life events considered to be particularly impactful events in the social life of an individual (eg, marriage, death of spouse or a close family member, pregnancy, change in residence, changes in working hours). Each life event has a prespecified weighting (ie, "life units") based on how traumatic the event felt to the large normative sample. The respondent indicates how many times each life event has occurred during the past 12 months or is expected to occur in the near future. The number of each life event is multiplied by the weights and summed to produce a total score of "life units."

Carer Activities

The questions about carer activities are adapted from the Nurse's Health Study II [23]. The respondents are asked whether they regularly provide care to a disabled or ill person, and, if applicable, how many hours per week they spend on such carer activities.

Social Media and Technology

In regard to the social media and technology domain, the questions are adapted from the Australian 2017 Sensis Social Media Report [27]. The questions focus on the respondent's use of the internet, use of social networking sites (eg, Facebook, Twitter, and Snapchat), type of devices used to access social networking sites, reasons for using social networking sites, and experiences with using social networking sites.

Health Record Linkage

Participants are asked to provide consent to have their personal health information retrieved from administrative data collections

(ie, ambulance, emergency department, hospital admissions, cancer registry, and mortality records) in New South Wales from January 1, 2017 to 12 months after the baseline survey. Secure data linkage will be conducted by the Centre for Health Record Linkage (CHeReL).

Data Management

All study information will be obtained, stored, and analyzed in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Research Involving Humans [28]. All results will be published in a form that will not allow any individual participants to be identified (ie, in tabular, aggregate form only). The participant registration form contains personal data (eg, name, residential address, email address, mobile phone number, and a relative's contact details). A participant ID number will be generated for all participants and stored with the data. The participant registration details will be stored separately in a secure password-protected folder. The data collected from the baseline and follow-up surveys will not contain any personally identifiable information, only the participant ID number to allow survey responses to be linked.

For participants who provide consent to have their health records linked to their survey responses, their personal data and participant ID number will be securely transferred to the CHeReL for the purposes of health record linkage. During the record linkage process, the CHeReL will generate a project person number (PPN) for each participant. The CHeReL will not have access to any of the collected data (ie, survey responses or health records). The PPNs are then linked to the existing participant ID numbers in each administrative data collection and returned to the respective data custodians. In turn, the data custodians for each administrative data collection (ie, ambulance, emergency department, hospital admissions, cancer registry, and mortality records) will securely transfer the health data records with PPNs to the researchers. Finally, the researchers will use the PPNs to link survey responses and health records belonging to the same individual, thereby creating a complete data set for analysis.

Data Analysis

Data will be analyzed using SAS version 9.4 (SAS Institute). The recruitment rate will be calculated as the number of registered participants divided by the total number of Macquarie University graduates in 2018. The denominator data will be supplied by the Macquarie University Graduation Office. The opt-in rates for data linkage of health records and survey responses will be calculated as the number of registered participants opting in divided by the number of registered participants. The representativeness of the sample will be evaluated by comparing its demographic profile with that of the full graduating cohort. Survey drop-out rates will be calculated separately for the baseline and follow-up surveys as the number of participants completing the survey divided by the number of participants starting the survey. Descriptive statistics will be used to evaluate the completeness of the baseline and follow-up surveys. The retention rate of the sample will be calculated as the number of participants completing the follow-up survey divided by the number of participants completing the baseline survey.

Results

Participant recruitment was completed in October 2018, and data collection for the baseline and follow-up surveys was completed in November 2019. As of April 2020, the process of acquiring health records from administrative data collections has commenced. The findings of this pilot cohort study will be prepared for publication in mid-2020. These findings will include the opt-in and opt-out rates for data linkage of health records and survey responses; a description of the representativeness of recruited participants; a description of the completeness of baseline and follow-up online survey items; and attrition rates for the 12-month follow-up survey.

Discussion

Emerging adulthood is a unique segment of an individual's life course. The defining features of this transitional period include identity exploration, instability, possibilities, self-focus, and feeling in-between adolescence and adulthood, all of which are thought to impact quality of life, health, and well-being. A longitudinal cohort study with a comprehensive set of measures would facilitate greater understanding of the multi-faceted elements and unique challenges that contribute to the health and well-being of emerging adults. Before expending significant resources on a large, longitudinal cohort study, it is advisable to first test the feasibility and inform the development of the larger study.

This pilot cohort study aims to evaluate the feasibility of recruiting Australian university graduates to establish a longitudinal cohort study to inform our understanding of emerging adulthood. It will evaluate the ability to recruit university graduates and obtain good quality survey data on a wide range of relevant measures. It is vital to obtain estimates of recruitment, response, and retention (or attrition) rates because these will inform the sample size and statistical power calculations that are necessary for planning and designing a future longitudinal cohort study. An evaluation of the measures in the pilot study is also necessary for the development and selection of the measures to be included in the main study.

There are challenges with recruiting university graduates into cohort studies using web-based surveys. Selection bias is an important consideration as many studies conducted among students report response rates below 20% [29-31]. University students and graduates frequently receive requests to participate in surveys, and this over-surveying can potentially lead to survey fatigue and poor response rates [32]. Compounding the issue is that the average response rate of web surveys is approximately 10% lower than that of mail or telephone surveys [33]. Student engagement, lottery incentives, and extra reminders can be effective for increasing the overall response rate [32,34-36]. However, merely increasing response rate does not necessarily entail diversifying or improving the representativeness of the sample [32]. Although self-selection can lead to unreliable survey outcomes [37], there are potential methods for correcting for selection biases (eg, poststratification or weighting class adjustments, propensity score adjustments, and generalized regression modelling [38]). It has been suggested that student

surveys with a 10% or lower response rate can eventually be considered trustworthy if the response quality is good [34].

Response bias is another pervasive problem in the design of surveys. The many types of bias are question design (eg, problems with wording, leading questions, faulty scales, intrusiveness), questionnaire structure (eg, formatting, priming, length, response fatigue), administration of questionnaire (eg, respondent's learning, recall, primacy or recency depending on mode) [39,40]. Some of these biases can be minimized by adopting previously validated instruments and scales. In the present study, the baseline and follow-up surveys are comprised almost entirely of commonly used and previously validated instruments and scales.

As to limitations, the study is modest in scale, and will be conducted at a single Australian university. This may not be representative of the broader population of university graduates in Australia.

In conclusion, this pilot study comes at a crucial time for research of this kind. It is expected that the findings from the pilot will identify areas for improvement and inform the development of a future longitudinal cohort study.

Acknowledgments

The study was funded by a Macquarie University Research Seeding Grant (grant reference number: 38278336). The funding body had no role in the design of this study, data analysis, interpretation of findings, or preparation of manuscripts. Ethical approval for this pilot cohort study has been granted by the Macquarie University Human Research Ethics Committee (reference number: 5201600265).

Authors' Contributions

All authors contributed to the conception and design of the study. DFP and RPL were responsible for drafting the first version of the manuscript. JA, JB, CMC, RJM, and VW were responsible for supervising the project. All authors contributed to or edited the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000;55(5):469-480. [doi: 10.1037/0003-066X.55.5.469] [Medline: 10842426]
2. Arnett JJ. Emerging adulthood: The winding road from the late teens through the twenties (2nd edition). New York, NY: Oxford University Press; 2014.
3. Buhl HM. Well-being and the child–parent relationship at the transition from university to work life. J Adolesc Res 2007;22(5):550-571. [doi: 10.1177/0743558407305415]
4. Tanner JL. Mental health in emerging adulthood. In: Arnett JJ, editor. The Oxford Handbook of Emerging Adulthood. New York, NY: Oxford University Press; 2015.
5. Barlett CP, Barlett ND. The young and the restless: Examining the relationships between age, emerging adulthood variables, and the Dark Triad. Pers Individ Dif 2015;86:20-24. [doi: 10.1016/j.paid.2015.05.024]
6. Barlett CP, Barlett ND, Chalk HM. Transitioning through emerging adulthood and physical health implications. Emerg Adulthood 2018 Dec 23. [doi: 10.1177/2167696818814642]
7. Centers for Disease Control and Prevention. Health-related quality of life (HRQOL): HRQOL concepts URL: https://www.cdc.gov/hrqol/concept.htm [accessed 2019-11-21]
8. Karimi M, Brazier J. Health, health-related quality of life, and quality of life: What is the difference? Pharmacoeconomics 2016;34(7):645-649. [doi: 10.1007/s40273-016-0389-9] [Medline: 26892973]
9. The World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization. Social Science & Medicine 1995 Nov;41(10):1403-1409. [doi: 10.1016/0277-9536(95)00112-k] [Medline: 8560308]
10. Pons-Villanueva J, Rodríguez de Armenta MJ, Martínez-González MA, Seguí-Gómez M. Longitudinal assessment of quality of life and its change in relation to motor vehicle crashes: The SUN (Seguimiento Universidad de Navarra) Cohort. J Trauma 2011;70(5):1072-1077. [doi: 10.1097/TA.0b013e3181eaad92] [Medline: 21131856]
11. van Oostrom SH, Smit HA, Wendel-Vos GCW, Visser M, Verschuren WMM, Picavet HSJ. Adopting an active lifestyle during adulthood and health-related quality of life: The Doetinchem Cohort Study. Am J Public Health 2012;102(11):e62-e68. [doi: 10.2105/AJPH.2012.301008] [Medline: 22994283]
12. NSW Health. NSW Population Health Survey Questionnaire 2016 URL: https://www.health.nsw.gov.au/surveys/adult/Documents/questionnaire-2016.pdf [accessed 2019-11-21]
13. Fair Work Commission. Australian Workplace Relations Study 2013-2014 Employee Relations (HR) Questionnaire URL: https://www.fwc.gov.au/documents/awrs/awrs%20employee%20relations%20questionnaire.pdf [accessed 2019-11-21]

14. Marks SR, MacDermid SM. Multiple roles and the self: A theory of role balance. J Marriage Fam 1996;58(2):417-432. [doi: 10.2307/353506]
15. DeCastro R, Griffith KA, Ubel PA, Stewart A, Jagsi R. Mentoring and the career satisfaction of male and female academic medical faculty. Acad Med 2014;89(2):301-311. [doi: 10.1097/ACM.0000000000000109] [Medline: 24362376]
16. Ware J, Kosinski M, Keller SD. A 12-item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. Med Care 1996;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
17. Buchholz I, Janssen MF, Kohlmann T, Feng Y. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. Pharmacoeconomics 2018;36(6):645-661. [doi: 10.1007/s40273-018-0642-5] [Medline: 29572719]
18. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
19. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
20. Peters L, Sunderland M, Andrews G, Rapee RM, Mattick RP. Development of a short form Social Interaction Anxiety (SIAS) and Social Phobia Scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. Psychol Assess 2012;24(1):66-76. [doi: 10.1037/a0024544] [Medline: 21744971]
21. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
22. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). Aust N Z J Public Health 2001;25(6):494-497. [doi: 10.1111/j.1467-842x.2001.tb00310.x] [Medline: 11824981]
23. Bao Y, Bertoia ML, Lenart EB, Stampfer MJ, Willett WC, Speizer FE, et al. Origin, methods, and evolution of the three Nurses' Health Studies. Am J Public Health 2016;106(9):1573-1581. [doi: 10.2105/AJPH.2016.303338] [Medline: 27459450]
24. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: Assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]
25. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The Multidimensional Scale of Perceived Social Support. J Pers Assess 1988;52(1):30-41. [doi: 10.1207/s15327752jpa5201_2]
26. Holmes TH, Rahe RH. The Social Readjustment Rating Scale. J Psychosom Res 1967;11(2):213-218. [doi: 10.1016/0022-3999(67)90010-4] [Medline: 6059863]
27. Sensis. Sensis Social Media Report 2017 URL: https://irp-cdn.multiscreensite.com/535ef142/files/uploaded/Sensis-Social-Media-Report-2017.pdf [accessed 2019-11-21]
28. National Health and Medical Research Council. National Statement on Ethical Conduct in Research Involving Humans 2007 (Updated 2018) URL: https://tinyurl.com/y679fhc3 [accessed 2019-11-21]
29. Lee JJ. International students' experiences and attitudes at a US host institution: Self-reports and future recommendations. J Res Int Educ 2010;9(1):66-84. [doi: 10.1177/1475240909356382]
30. Sax L, Gilmartin S, Bryant A. Assessing response rates and nonresponse bias in web and paper surveys. Res High Educ 2003;44(4):409-432. [doi: 10.1023/A:1024232915870]
31. Lauber C, Ajdacic-Gross V, Fritschi N, Stulz N, Rössler W. Mental health literacy in an educational elite -- An online survey among university students. BMC Public Health 2005;5:44. [doi: 10.1186/1471-2458-5-44] [Medline: 15882465]
32. Van Mol C. Improving web survey efficiency: The impact of an extra reminder and reminder content on web survey response. Int J Soc Res Methodol 2016;20(4):317-327. [doi: 10.1080/13645579.2016.1185255]
33. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Comput Human Behav 2010;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]
34. Nair CS, Adams P, Mertova P. Student engagement: The key to improving survey response rates. Qual High Educ 2008;14(3):225-232. [doi: 10.1080/13538320802507505]
35. Laguilles JS, Williams EA, Saunders DB. Can lottery incentives boost web survey response rates? Findings from four experiments. Res High Educ 2010;52(5):537-553. [doi: 10.1007/s11162-010-9203-2]
36. Saleh A, Bista K. Examining factors impacting online survey response rates in educational research: Perceptions of graduate students. J Multidiscip Eval 2017;13(29):63-74 [FREE Full text]
37. Bethlehem J. Selection bias in web surveys. Int Stat Rev 2010;78(2):161-188. [doi: 10.1111/j.1751-5823.2010.00112.x]
38. Greenacre ZA. The importance of selection bias in internet surveys. Open J Stat 2016;6(3):397-404. [doi: 10.4236/ojs.2016.63035]
39. Choi BC, Pak AW. A catalog of biases in questionnaires. Prev Chronic Dis 2005;2(1):A13. [Medline: 15670466]
40. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. J Appl Psychol 2003;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]

Abbreviations

ACTRN: Australian New Zealand Clinical Trials Registry
BRS: Brief Resilience Scale
CHeReL: Centre for Health Record Linkage
EQ-5D: EuroQoL 5-dimension
GAD-7: General Anxiety Disorder scale
HRQOL: health-related quality of life
K10: Kessler Psychological Distress Scale
MQ-MINDS: Macquarie University – Monitoring of Injury and Psychosocial Health Outcomes, Career Trajectories and Continuing Education, Lived Experiences, and Social Connectedness
MSPSS: Multidimensional Scale of Perceived Social Support
RBS: Role Balance Scale
SF-12: Short Form Health Survey
SIAS-6: Social Interaction Anxiety Scale
SRRS: Social Readjustment Rating Scale

Edited by G Eysenbach; submitted 03.09.19; peer-reviewed by G Hu, S Badawy, T Hendriks, F Gomez; comments to author 11.11.19; revised version received 02.12.19; accepted 17.03.20; published 23.04.20.

Please cite as:
Lystad RP, Fajardo Pulido D, Peters L, Johnstone M, Ellis LA, Braithwaite J, Wuthrich V, Amin J, Cameron CM, Mitchell RJ
Monitoring Health and Well-Being in Emerging Adults: Protocol for a Pilot Longitudinal Cohort Study
JMIR Res Protoc 2020;9(4):e16108
URL: http://www.researchprotocols.org/2020/4/e16108/
doi:10.2196/16108
PMID:32324142

Protocol

Biopsychosocial Mechanisms Linking Gender Minority Stress to HIV Comorbidities Among Black and Latina Transgender Women (LITE Plus): Protocol for a Mixed Methods Longitudinal Study

Ashleigh J Rich[1], MPH; Jennifer Williams[2], PhD; Mannat Malik[3], MHS; Andrea Wirtz[3], MHS, PhD; Sari Reisner[4], MA, SCD; L Zachary DuBois[5], PhD; Robert Paul Juster[6], PhD; Catherine R Lesko[3], MPH, PhD; Nicole Davis[2], MPH; Keri N Althoff[3], PhD; Christopher Cannon[7], MPH; Kenneth Mayer[8], MD; Ayana Elliott[9], DNP; Tonia Poteat[2], MPH, PhD

[1]School of Population & Public Health, University of British Columbia, Vancouver, BC, Canada

[2]Department of Social Medicine, University of North Carolina, Chapel Hill, Chapel Hill, NC, United States

[3]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[4]Harvard Medical School, Harvard University, Boston, MA, United States

[5]Department of Anthropology, University of Oregon, Eugene, OR, United States

[6]Department of Psychiatry and Addiction, University of Montreal, Montreal, QC, Canada

[7]Whitman Walker Health, Washington, DC, DC, United States

[8]TH Chan School of Public Health, Harvard University, Boston, MA, United States

[9]National LGBT Health Education Center, Boston, MA, United States

Corresponding Author:
Ashleigh J Rich, MPH
School of Population & Public Health
University of British Columbia
2206 East Mall
Vancouver, BC, V6T 1Z8
Canada
Phone: 1 6043637224
Email: ajrich@mail.ubc.ca

Abstract

Background: Black and Latina transgender women (TW) experience a disparate burden of HIV and related comorbidities, including poor mental health and cardiovascular disease (CVD) risks. Pervasive multilevel stigma and discrimination operate as psychosocial stressors for TW living with HIV and shape health disparities for this population. Gender-affirming hormone therapy (GAHT) is commonly used by TW to facilitate alignment of the body with gender identity; in the context of stigma, GAHT may both improve mental health and increase CVD risks.

Objective: This study aims to quantify the longitudinal relationship between stigma and chronic stress among black and Latina TW living with HIV. Secondary objectives include identifying pathways linking chronic stress to HIV comorbidities and exploring chronic stress as a mediator in the pathway linking stigma and GAHT to CVD comorbidities.

Methods: This US-based mixed methods longitudinal study will enroll a prospective cohort of 200 black and Latina TW living with HIV, collecting quantitative survey data, qualitative interviews, and biomarkers of chronic stress. Interviewer-administered surveys will include validated psychosocial measures of self-reported stigma and discrimination, perceived stress, CVD risk factors, mental health, access to gender-affirming care, coping, and social support. Medical record abstraction will collect data on GAHT use, CD4 count, HIV viral load, antiretroviral therapy, treatment, and comorbid conditions. Clinical measures will include physiological biomarkers as well as salivary and blood-based biomarkers of chronic stress. Survey data will be collected every 6 months (baseline, and 6, 12, 18, and 24 months), and biospecimens will be collected at baseline and at 12 and 24 months. A purposive subsample (stratified by use of GAHT and presence of depressive symptoms) of 20 to 30 TW living with HIV will be invited to participate in in-depth interviews at 6 and 18 months to explore experiences of intersectional stigma, chronic stress, and the role of GAHT in their lives.

Results: This study was funded by the National Institute on Minority Health and Health Disparities in December 2018. The study community advisory board and scientific advisors provided critical input on study design. Recruitment began in October 2019 (n=29 participants as of submission) and data collection will continue through 2022, with publication of baseline results anticipated summer 2021.

Conclusions: This study will focus on black and Latina TW living with HIV, an understudied health disparities population, advance both stigma and intersectionality research, and move chronic stress physiology research toward a more nuanced understanding of sex and gender. The comprehensive methodology will support the exploration of the role of exogenous estrogen in the pathways between stress and HIV comorbidities, elucidating the role of GAHT in the stress-health relationship. Finally, this study will provide longitudinal evidence of the impact of stigma-related chronic stress on the lives of black and Latina TW living with HIV integrating qualitative and quantitative data with psychosocial, clinical, and biological measures.

International Registered Report Identifier (IRRID): DERR1-10.2196/17076

(JMIR Res Protoc 2020;9(4):e17076) doi:10.2196/17076

KEYWORDS

transgender persons; HIV; comorbidity; racial factors; stress, physiological; stress, biological

Introduction

HIV Disparities Among Transgender Women

Transgender women (TW) experience a disparate burden of HIV infection and HIV-related comorbidities [1]. A recent systematic review and meta-analysis of HIV among transgender populations in the United States found that HIV prevalence ranged from 14.2% by laboratory-confirmed diagnosis to 21.0% by self-report for TW [2]. Globally, overall HIV prevalence for TW is estimated to be 19%, 49-fold higher odds compared with that among cisgender adults [3]. Among TW, the heaviest burden of HIV is borne by black and Latina TW (BLTW) [4,5], with an overall HIV prevalence of 44.2% for black TW and 25.8% for Latina TW [2]. BLTW make up the majority of TW receiving HIV clinical care nationally in the United States [6,7].

Multiple Pathways to Health Disparities for Transgender Women

Multilevel and multidimensional factors shape health disparities among TW living with HIV (TWLHIV). Individual (eg, sociodemographic and psychological), interpersonal (eg, violence, victimization, and gender-power dynamics), and structural (eg, stigma and discrimination) factors influence HIV-related outcomes among TWLHIV [4,8-11]. TWLHIV have high rates of living in poverty, facing housing insecurity, and lacking health insurance [7]. Given these barriers and experiences that fundamentally undermine health and well-being, TWLHIV are also less likely to adhere to antiretroviral therapy (ART) or to achieve durable HIV viral load suppression [7], thus facing elevated risk of mortality [12].

Due in part to stigma exposure, mental health and cardiovascular disease (CVD) disparities are common synergistic HIV comorbid conditions experienced by TWLHIV. In the largest national survey of transgender people to date (N=27,715), 40% reported ever attempting suicide and 39% reported psychological distress in the prior year, compared with 4.6% and 5% of the US general population, respectively [13]. A growing body of literature has identified associations between poor mental health and exposure to transgender stigma [14-17]. This may be exacerbated by HIV infection, as depression is the most common neuropsychiatric comorbidity among people living with HIV (PLWHIV), associated with poor adherence, lack of viral suppression, and increased CVD risk [18-20].

TW are also more likely to experience CVD risk factors, events, and mortality than cisgender adults (ie, nontransgender adults) [21,22]. In one of the largest studies published to date on TW and CVD, a retrospective mortality analysis of more than 1000 Dutch transgender adults, a 64% increased risk in CVD mortality was observed among TW compared with the general population [21]. Increased prevalence of CVD among PLWHIV has been attributed to chronic inflammation associated with HIV infection as well as higher prevalence of CVD risk factors (eg, obesity and diabetes) [22], risk behaviors (eg, smoking) [23,24], and stigma [25]. However, underlying psychologic and biologic pathways to CVD disparities among TWLHIV are poorly understood [26].

Associations of Gender-Affirming Hormone Therapy With Chronic Stress and Cardiovascular Disease

Gender-affirming hormone therapy (GAHT), including exogenous estrogen, is commonly used by TW to facilitate alignment of the physical body with gender identity [13]. Approximately 75% to 95% of TW take GAHT at any point in time [27-29]. Access to GAHT is a community priority for TW [13] and can be a critical protective factor for HIV and other comorbidities [30], improving psychological functioning [21,31], facilitating care engagement [32], and improving ART adherence and viral suppression when provided in the context of HIV care [31,33-35]. In contrast to these benefits, GAHT has also been associated with an elevated CVD risk [21,22,27,36] and may potentiate CVD comorbidities among TWLHIV. Sex hormones play an important modulatory role in stress physiology [37,38]. They have been implicated in sex differences in CVD risk [39] and mental health [40]. However, previous studies in this area were conducted with cisgender people only or did not consider gender experience, limiting the ability to disaggregate hormonal effects from other gendered factors. Clinicians, scientists, and transgender communities have called for more longitudinal research on the effects of GAHT on health outcomes among TWLHIV [9,11], and especially among older TW who have been particularly understudied [36]. Specifically, data are needed on how GAHT

may impact health disparity pathways for TWLHIV, who face both mental health and CVD comorbidities.

Allostatic Load as a Biological Marker of Chronic Stress

Allostatic load (AL) refers to the cumulative *wear and tear* effects of chronic stress on the brain and body [41,42]. Models of AL demonstrate physiologic pathways linking psychosocial stressors to poorer physical and mental health [43]. AL derives from the concept of allostasis [44], the dynamic adaptive regulatory process of the body that seeks to maintain homeostasis during exposure to physical and psychological stressors [45]. As such, AL can be measured by assessing neuroendocrine, immune, metabolic, and cardiovascular biomarkers [45-47]. Dysregulation of these biomarkers has been linked to minority stress experiences, including stigma and discrimination, and is associated with poor mental health and CVD among racial and ethnic minorities [46-49], as well as elevated CVD risk among sexual minorities [50]. AL differs by sexual orientation [51]; however, studies examining pathways among transgender people who experience gender minority stress is limited [52,53], and there are few published studies of physiologic stress processes among PLWHIV [54]. Mechanisms linking chronic stress to comorbidities have not been elucidated for transgender people, particularly TW who experience stigma and discrimination on the basis of intersecting minority identities, such as race, gender identity, and HIV status (ie, intersectional stigma) [55,56]. Qualitative research among black women living with HIV has highlighted the inextricability of gender and race in understanding HIV-related stigma and its sequelae [57].

Project Proposal

This novel prospective study aims to advance scientific knowledge of how intersectional stigma impacts HIV outcomes and comorbidities for black and Latina TWLHIV. Elucidating the multilevel pathways linking intersectional stigma to mental health and CVD comorbidities among TWLHIV, and exploring the role GAHT may play in mitigating or exacerbating these comorbidities in the context of gender minority stress, will advance scientific understanding of how stigma and discrimination become embodied. These advances in our understanding are key to our ability to identify ways clinical providers can more effectively tailor their care to meet the needs of this health disparity population. This project employs a mixed methods protocol in which biomarkers, clinical measures, survey responses, and in-depth interviews will be used to understand pathways shaping HIV comorbidities for black and Latina TWLHIV. Results will inform clinical practices and public health interventions that facilitate health care engagement and retention to promote health equity among black and Latina TWLHIV.

Objectives

The primary objectives of this study are as follows:

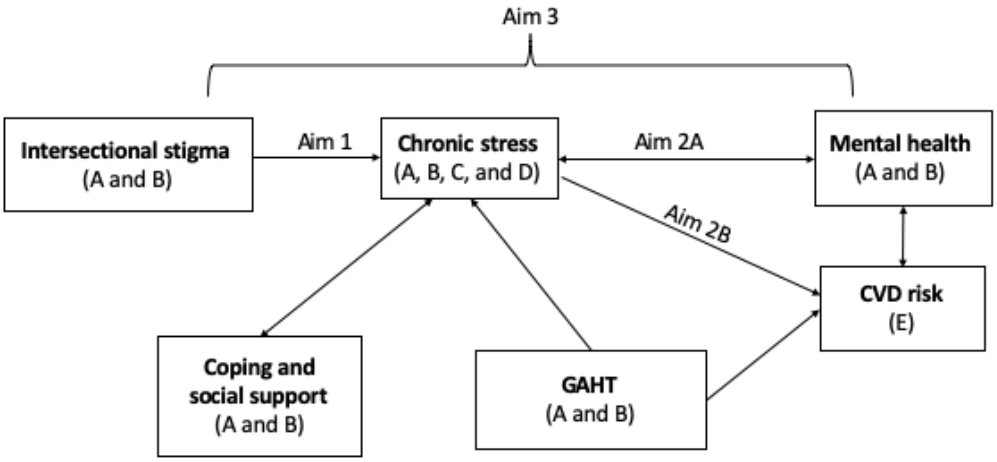
1. To quantify the longitudinal relationship of stigma to chronic stress biomarkers in TWLHIV;
2. To identify pathways linking chronic stress biomarkers to comorbidities (ie, CVD and mental health) among TWLHIV; and
3. To explore chronic stress as a mediator in pathways linking stigma, GAHT, and HIV comorbidities.

Methods

Conceptual Model

This study, named LITE Plus, draws on the Gender Minority Stress and Resilience Model (GMSR) [58,59] and chronic stress and AL constructs [60]. Our research uses an intersectionality [56] lens to recognize the multiple and intersecting levels of influence that shape health disparities among transgender populations. The GMSR posits that transgender mental and physical health disparities are driven by distal (eg, discrimination) and proximal (eg, internalized stigma) stressors that can be mitigated by factors such as social support (eg, community connectedness) [58]. AL models [61] allow the objective measurement of multisystem physiologic dysregulation caused by chronic stress [43] and have been applied to the study of a variety of health disparities [62]. An intersectionality approach centers the experience of people from multiply marginalized groups, such as black and Latina TWLHIV, and acknowledges that these multiple social identities are associated with disparate health outcomes via a confluence of social forces, such as racism, misogyny, transphobia, and HIV stigma [55]. In centering the experiences of historically marginalized groups, an intersectionality lens encourages the examination of oppressive social processes without the need for comparison with dominant groups [63]. We will examine biological pathways linking the chronic stress of intersectional stigma to mental health and CVD comorbidities among TWLHIV and the potential mitigating and exacerbating influences of social support and coping and GAHT along these pathways (Figure 1**)**.

Figure 1. Hypothesized pathways linking stigma, stress, mental health, and cardiovascular disease outcomes among transgender women living with HIV to be measured using (A) survey, (B) qualitative interviews, (C) salivary cortisol, (D) allostatic load index, and (E) clinical measures. GAHT: gender-affirming hormone therapy; CVD: cardiovascular disease.



Study Design and Population

This study builds on the American Cohort to Study HIV Acquisition among Transgender Women in High Risk Areas (known as the LITE study) [64], which prospectively follows TW who are not living with HIV. Participants for the LITE Plus study will be recruited from two LITE clinical study sites (Fenway Health in Boston, MA, and Whitman-Walker Health in Washington, DC) that provide comprehensive health services to the lesbian, gay, bisexual, transgender, and queer communities, including HIV treatment and GAHT. Fenway consists of six satellite sites and Whitman-Walker consists of three health center sites. This mixed methods, multisite longitudinal study will enroll a prospective cohort of 200 black and Latina TWLHIV living in Boston, MA, and Washington, DC, areas (approximately 100 participants per site) and collect data every 6 months for 24 months total—at baseline and at 6, 12, 18, and 24 months. Interview-administered surveys will include validated psychosocial measures of stigma, perceived stress, CVD risk behaviors (eg, smoking), mental health, coping, and social support. Medical record abstraction will collect data on GAHT use, CD4 count, HIV viral load, ART medications, and comorbid conditions. Clinical measures will include height, weight, waist circumference, and blood pressure; salivary measures of cortisol; and blood samples for measures of physiological stress to be included in the AL index. Qualitative in-depth interviews will be conducted with a subset of 20 to 30 TWLHIV (up to 15 participants from each site, and based on participants' use of GAHT and risk for depression) to explore their experiences of intersectional stigma, chronic stress, and the role of GAHT in their lives.

A community advisory board (CAB) has reviewed the study design and implementation plans to ensure that retention, recruitment, and data collection align with gender-affirming best practices. The LITE Plus CAB meets annually to receive study updates and give feedback on study progress. They are also consulted between meetings for feedback on study forms, methods, and surveys. The CAB will also review preliminary results and offer interpretations of findings as part of a community-engaged research model.

Inclusion Criteria

Participants must have been assigned male sex at birth; identify as *female*, *woman*, *trans female*, *male-to-female*, or *woman of trans experience*; have a laboratory-confirmed HIV diagnosis; identify as black, Latina, and/or multiracial (inclusive of black and/or Latina identities); and be 18 years of age or older. All study participants must live in or near Boston, MA, or Washington, DC, speak English and/or Spanish, be mentally sound and capable of consenting and provide written consent to participate. To be eligible for home collection of salivary cortisol, participants must not be currently using any medicines containing steroids regardless of route of administration, have consistent access to a freezer for up to 4 weeks at a time, live within established geographic boundaries based on participant zip codes to facilitate staff retrieval of samples, and report willingness to collect three sets of six saliva samples over the course of the study. Eligibility for the qualitative interviews will be based on the participant history of GAHT and depression scores at baseline study visit.

Study Procedures

Recruitment

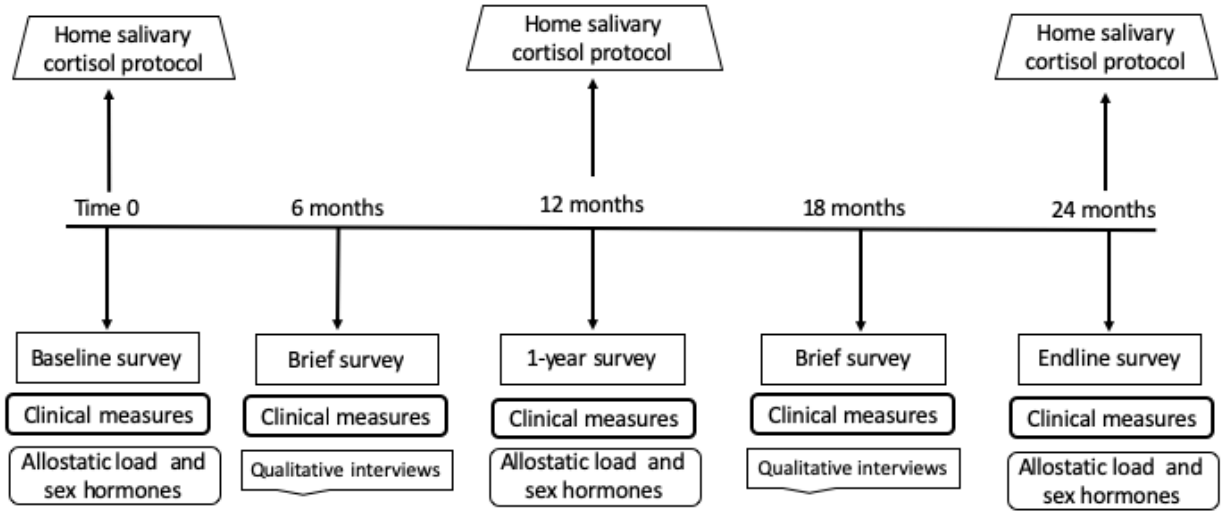
Multiple strategies will be used to identify, recruit, and retain participants. Participants may be identified as eligible via their electronic medical records and invited to participate by study staff. Participants may also be recruited from the LITE study, when (1) they are excluded from the cohort at baseline because of a positive HIV test result, or (2) when they seroconvert during the follow-up, which ends their LITE cohort participation. Should we not reach target enrollment through clinic-based recruitment and recruitment through LITE, we will engage in community-based recruitment, in partnership with LITE. Community-based strategies will include venue-based recruitment from the gender-affirming community events and organizations frequented by TW and social media outreach.

Visit Schedule

Surveys, clinical measures, and medical record abstraction will be collected at all study visits (every 6 months) over the 24-month study period. Phlebotomy and salivary cortisol

self-collection training will occur at the baseline, 12-month, and 24-month visits. At the 6- and 18-month visits, a subgroup of participants will be invited to participate in in-depth interviews. The data collection schedule is detailed in Figure 2.

Figure 2. Overview of study visits and data collection intervals.



Survey

Each participant will complete an interviewer-administered survey at all study visits over the 24-month study period. Participants will complete a longer survey at the baseline, 12-month, and 24-month visits, with brief surveys at the 6- and 18-month visits. Surveys will be administered by a trained interviewer via a tablet or desktop computer and will include validated measures used in prior research with transgender communities, where available. Survey domains include sociodemographics, validated psychosocial measures of stigma and discrimination, perceived stress, CVD risk behaviors, mental health, resilient coping, material social support, community connectedness, interpersonal violence, engagement in HIV care, and self-reported GAHT use. Baseline survey measures are described in Table 1.

Table 1. Key quantitative measures of the baseline survey.

Construct	Measures
ART[a] adherence	Self-reported use, treatment interruptions, recent missed doses and reasons for missed doses [65], and challenges obtaining ART
Chronic stress	Perceived stress [66]
Coping and resilience	Brief resilient coping [67], gender identity pride [58], and community connectedness [58]
GAHT[b]	Duration, source, adherence, mode of delivery, frequency of monitoring of hormone levels in blood, discussion with health care provider regarding potential side effects, and perceptions of GAHT and ART interaction
Gender-affirming surgery	History of and need for
Health care access	Health insurance, health care access barriers, and typical health care setting
General health	History of diabetes, hypertension, high cholesterol, heart disease, blood clots, stroke, kidney disease, liver disease, cancer, obesity, etc, and perceived general health and healthy days (HRQOL-4[c]) [68]
HIV outcomes	Last time viral load measured, if applicable: reasons for not having viral load measured recently (within 6 months), suppressed viral load at last measurement
Intersectional stigma and discrimination	Fear of deportation [69,70], internalized HIV stigma (within social relationships) [71], internalized anticipated discrimination [72], everyday discrimination [73], internalized transphobia [58], and gender-related rejection [58]
Legal gender transition	Congruence between gender and preferred name and gender marker and name listed on IDs and records, and importance of congruent IDs and records [74]
Medical distrust	Trust in HIV care providers [75]
Mental health	Posttraumatic stress disorder (PCL-C[d]) [76], depressive symptomology (CESD-10[e]) [77], and history of suicidality and attempted suicide
STI[f]	History of STI testing and diagnosis, and history of hepatitis C
Sex work	Lifetime and recent history of engagement in sex work
Smoking history	Current smoking status [78] and pack-year smoking history [79]
Social support	Material social support [80]
Sociodemographics	Gender identity, sexual orientation, completed education, employment status, housing status and homelessness, immigration status, and material hardship [81]
Soft-tissue fillers	Lifetime use, location (ie, body parts), source of injections (eg, medical provider and parties)
Substance use	Past-year alcohol use (AUDIT-C[g]) [82] and past-year drug use (DAST-10[h]) [83]
Violence experiences	Lifetime and recent psychological, physical, and sexual violence (RCTS-2[i]) [84]

[a]ART: antiretroviral therapy.

[b]GAHT: gender-affirming hormone therapy.

[c]HROOL-4: Healthy Days Core Module health related quality of life measure.

[d]PCL-C: Post-traumatic stress disorder checklist- civilian version.

[e]CESD-10: Centre for Epidemiological Studies Depression Scale.

[f]STI: sexually transmitted infection.

[g]AUDIT-C: Alcohol Use Disorders Identification Test-Alcohol Consumption Questions.

[h]DAST-10: Drug Abuse Screening Test.

[i]RCTS-2: Revised Conflict Tactics Scales.

Clinical Measures and Biospecimens

Clinical measures will be collected at each study visit, specifically height (at baseline only), weight, waist circumference, and blood pressure. Biospecimens (blood and saliva) will be collected at baseline and at 12- and 24-month follow-up visits to test for chronic stress biomarkers and CVD risk. At-home saliva collection will occur following the baseline, 12- and 24-month study visits. We chose to use saliva to test for stress-related biomarkers because at-home collection of saliva is considered minimally invasive [85]. Home salivary collection has proven feasible and effective in previous studies of stress and health among transgender men [37]. Furthermore, prior studies of at-risk populations (eg, abused women and low-income populations) have successfully utilized salivary analytes to study stress [86,87]. Salivary cortisol has been used extensively as a biomarker of stress in research settings, especially in studies examining psychological stress with repeated measurements [88], and is a useful component of the Allostatic Load Index (ALI) [89]. As cortisol follows a diurnal

pattern throughout the day (peaking within 30 min of waking and declining over the course of the day), each participant will have a unique diurnal curve measured as awakening, 30-min postwake, and bedtime cortisol values [85,90,91]. Participants will be trained in the *passive drool* method of saliva self-collection [90] during baseline and receive booster trainings at the 12- and 24-month study visits. They will be provided with supplies to self-collect their saliva at home on two consecutive weekdays as soon as possible after the study visit. Participants will be instructed to store home-collected saliva samples in their freezers until study staff collect the samples, typically within 4 weeks. Samples will be subsequently stored in a −80 freezer at the site until they can be mailed to the designated laboratory for analysis. Sites will ship frozen saliva samples in batches to the Institute for Interdisciplinary Salivary Bioscience Research at the University of California Irvine every 6 months for analysis and storage.

In addition, study staff will draw 44 mL (3 tablespoons) of blood from each participant to test for estradiol, testosterone, progesterone, dehydroepiandrosterone-sulphate (DHEAS), total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein, triglycerides, interleukin-6 (IL-6), tumor necrosis factor-alpha (TNF-alpha), C-reactive protein (CRP), fibrinogen, insulin, glycosylated hemoglobin (Hb), albumin, and creatinine levels. Blood samples will be sent to the LabCorp location nearest to the participant's study site for analysis and results, once available, will be entered into the study database and shared with the participant (along with cortisol results) at their subsequent follow-up visits.

Concomitant Medications

Study staff will collect information on concomitant medication use from participants, supplemented by medical record review. This will include all medications taken within 30 days of study visits. Concomitant medications of interest include ART, GAHT, and medications for blood pressure, diabetes, and heart disease.

Medical Record Review

Medical record review and data extraction will be done for each participant within 1 week of each study visit. Participants will sign a Health Insurance Portability and Accountability Act (HIPAA) waiver as part of the study consent form. Participants that do not have medical records at Fenway or Whitman-Walker will sign a medical record request form that will be sent to their medical care provider. The medical record review and data extraction will be completed by trained data collectors using a standardized form. Specific information to be extracted from the medical records includes most recent CD4 count and HIV-RNA (viral load), current problem list and/or diagnosis codes, current medication list, and full history of all antiretroviral medications ever taken by the participant.

In-Depth Interviews

At 6 and 18 months, a stratified purposive subsample of 20 to 30 participants (up to 15 at each site) will be invited to participate in qualitative interviews. Participants will be stratified based on self-reported baseline GAHT use and dichotomized depression scores. Stratification by these parameters will allow for qualitative exploration of relationships between GAHT use and mental health, consistent with the third study aim to assess relationships among GAHT, chronic stress, and HIV comorbidities. Participants in the qualitative data collection will take part in two in-depth life history calendar (LHC) interviews, conducted virtually with study staff via HIPAA-compliant audiovisual communication software. LHCs are a form of participant-empowered data collection that allows participants to take an active role in the data collection process and give feedback on how data are documented while providing a descriptive timeline of life experiences [92]. LHCs have been successfully adapted for electronic use, including among TW [93], allowing participants and interviewers to complete calendars in a virtual format [94]. Using virtual (computer-based) LHC [95], TWLHIV will work with interviewers to create a longitudinal timeline of their gender journeys, mental health, experiences of stigma and minority stress and use of GAHT, anchored by significant life course milestones. These interviews will be conducted by data collectors with specialized training in qualitative data collection. Participants may use their personal computers or computers made available to them at the study site.

Incentives and Retention

Stepped incentives will be used to promote retention; participants must complete all study visit tasks to receive the incentive. Participants will be paid up to US $300 over the course of the study if they complete all study visit activities. Eligible participants may receive an additional US $150 for completing all of the home salivary collection and an additional US $60 for completing two in-depth interviews. An end-of-study bonus will be available for participants who are retained for 24 months and complete all study visits, making the incentive for the final visit US $80 for those participants. The detailed incentive structure is shown in Table 2.

The study clinical management system utilizes mobile phone text messaging and email to contact participants. The system automatically reminds participants of upcoming appointments and overdue visits, and sends other critical notifications regarding home salivary sample collection, distribution of incentives, etc. These notifications will serve as a first-line retention tool in combination with personalized messages and phone calls to encourage retention. Every effort will be made to coordinate study visits with routine clinical care visits for participants who also receive care at the study site.

All study information will be deidentified through the use of a unique identifier, which is generated for each participant at enrollment. Access to data by study staff will be on a role-based standard. All study staff will be trained in security and confidentiality procedures and will sign a confidentiality agreement before receiving access to any participant data. We will minimize the indirect disclosure of HIV status by referring to the study as a *health study for transgender women*.

Table 2. Stepped incentive structure.

Visit	Incentive amount (US $)
Baseline	50
6 months	50
12 months	60
18 months	60
24 months	70
If eligible	
Baseline saliva collection	50
12-month saliva collection	50
24-month saliva collection	50
In-depth Interview	30
Study completion bonus	10

Data Management and Tracking Participant Progress

LITE Plus uses the Clinical Trials Management System (CTMS) to track participant progress and automate study reminders. The automated features of the CTMS make it easy to remind participants of upcoming appointments and alert study staff when participant visit windows are closing. This HIPAA-compliant system provides a secure database in which to collect and store study data using the Transport Layer Security 2048-bit encryption. The CTMS is hosted by the Johns Hopkins University and has been customized to fit the LITE Plus protocol and workflow. Access to the CTMS is restricted to trained, certified data collectors and study team personnel.

Quality Assurance and Control

Data collectors have been trained in study-specific quality assurance guidelines that include checks for data quality and completion during and after participant visits. Data collectors are trained to check study forms for completion before marking forms as *completed* in the CTMS. A quality assurance and control tracker has been developed for each site. Trained data collectors use this tracker to ensure that study forms are complete and consent forms are signed and dated.

Statistical Methods and Analysis

Sample Size Calculation

Previous experience of this study team recruiting longitudinal cohorts of TWLHIV suggest 80% to 90% retention can be expected over the course of the study (ie, at the end of year 2, expected n=160) [96]. With a baseline sample of 200 TWLHIV, the power analysis shows a minimum detectable R^2 of 4.3%, which is between Cohen thresholds of 2% and 13% for small and medium minimum detectable effect sizes, respectively, assuming N=200 followed by 20% attrition to yield a minimum analysis N of 160 for the analyses proposed for aims 1 and 2 [97,98]. A sample size of 160 achieves 80% power to detect an R^2 of 0.043 attributed to one independent variable(s) using an *F* test with a significance level (alpha) of .050. The variables tested are adjusted for an additional 10 independent variable(s) with an R^2 of 0.090.

Quantitative Analysis

Univariate and bivariate analyses and multivariable regression models will be used to compare levels and sources of stress between different levels of reported intersectional stigma, accounting for individual differences (eg, age, CD4 count, and HIV-RNA). Multivariable analyses will be used to measure the effects of the independent variables (eg, intersectional stigma) on dependent variables (eg, AL) while holding constant some factors (eg, age). For longitudinal analyses, marginal structural models will be used to account for repeated measures with time-dependent confounding and the potential for confounders at one point to be mediators at another [99].

Qualitative Analysis

LHC interviews will be transcribed verbatim. Spanish language transcripts will then be translated. All transcripts and LHC visual data will be uploaded into Atlas.ti Scientific Software Development GmbH to facilitate analysis. Documents will be coded using a priori codes created based on the interview guides. Codebooks will be modified iteratively based on emergent themes as coding progresses. Two coders will analyze the same transcripts separately. Any discrepancies will be discussed and resolved by peer debriefing and consensus with the PI, making final decisions should consensus not occur. The life course framework will be used to analyze codes for patterns and themes [100]. Member checking will be conducted with CAB members to enhance rigor of analyses.

Results

Recruitment to Date

The patient population at Whitman-Walker Health includes 1243 TW, of whom 235 are living with HIV. Of these TWLHIV, 71.0% (167/235) are black and 22.1% (52/235) are Latina. Based on a prior experience recruiting 112 black and Latina TWLHIV in the span of 3 months for a previous TW study at this site, the study team expects to recruit 100 TWLHIV from Whitman-Walker Health for this study without difficulty. Fenway's patient population includes over 3500 transgender individuals of whom 1575 are TW. Among TWLHIV at Fenway,

more than half are black and one-quarter are Latina. In 2017, up to 40 new transgender patients initiated care every month at Fenway, and the number of TWLHIV has steadily increased. In addition to existing patients, LITE Plus will also recruit potentially eligible TW from the LITE cohort who have been excluded at baseline for a positive HIV test result (n=65 to date) or who may seroconvert during the follow-up. If necessary, additional participants will be recruited using community-based strategies.

This study was funded in December 2018 by the National Institute on Minority Health and Health Disparities for a start date of December 1, 2018, and an end date of March 31, 2023. The study protocol was approved using a single institutional review at the University of North Carolina at Chapel Hill Institutional Review Board (IRB 18-2632). Recruitment began in October 2019, and we aim to enroll the full cohort (N=200) within 9 months of beginning the enrollment. Retention is expected to be 90% over the course of the first 12 months of the follow-up and at least 80% over the study duration (24 months). As of submission, 29 participants have been enrolled in the study. Publication of baseline results is expected in summer 2021.

Main Study Constructs

Collected data will be used to construct the main study outcomes and predictors mentioned below.

Allostatic Load Index

As a way to examine cumulative stress effects on multiple body systems, AL is most often measured using an index that includes a battery of stress-related biomarkers with subclinical thresholds to quantify physiological dysregulations [101]. Over 100 studies have used a variety of AL algorithms that summarize neuroendocrine, immune, metabolic, and cardiovascular functioning [102] and predict disease better than existing approaches with single biomarkers [103,104]. Using 17 biomarkers (salivary cortisol, IL-6, TNF-alpha, DHEAS, insulin, glycosylated Hb, fibrinogen, CRP, total cholesterol, HDL, triglycerides, albumin, creatinine, systolic blood pressure, diastolic blood pressure, and BMI), we will calculate an ALI for each participant using an established, count-based approach that sums the number of dysregulated biomarkers using high-risk cutoffs based on the sample's distribution of values for each biomarker, that is, the 75th percentile for biomarkers for which high levels are harmful, or the 25th percentile for which low levels are harmful [102]. A review of nearly 60 empirical studies suggests that ALIs incorporating similar subclinical ranges for numerous biomarkers (mean=10, range=4-17) predict clinical outcomes better than methods that address only clinical thresholds [102]. As each biomarker is dichotomized as 0 or 1, each has an equal weight in calculating the ALI [102] and ALI scores can range from 0 to 17 for this study. Based on an early review, the biomarkers included in this approach represent those most commonly used in the ALI literature [102]. ALI will be calculated for each participant using clinical and laboratory measures collected at baseline, 12 months, and 24 months to allow for observation of change over time.

Gender-Affirming Hormone Use and Cardiovascular Disease Risk

Cumulative GAHT exposure will be estimated as self-reported total number of years taking GAHT at baseline. For example, a participant who is 40 years old and started taking GAHT at 20 years old without interruption will be categorized as having 20 years of GAHT exposure. The Atherosclerotic CVD Pooled Cohort Equations (PCE) estimator [105] will be used to calculate CVD risk. As scores on the PCE are calibrated by sex, we will use a threshold of 1 year of current, continuous GAHT exposure at baseline as the cutoff for using the female-calibrated PCE, that is, participants with 1 year of GAHT will be classified as female for the PCE while those with less than 1 year of GAHT will use the male-calibrated PCE. We will assess the relationship between baseline ALI scores and changes in the PCE scores during the study follow-up period. PCE scores at 24 months will be used for the CVD risk outcome. Baseline GAHT exposure will be used to calibrate the PCE in terms of sex, to ensure any changes in the PCE risk over the study period are because of changes in the other biomarkers used to calculate the PCE.

Mental Health Comorbidity

Participants will complete brief validated mental health and substance use screeners during study visits, which have been implemented in previous research with transgender people [106]. The PCL-C Checklist-Civilian [76], a continuous measure of posttraumatic stress disorder symptom severity, and the Centre for Epidemiological Studies Depression Scale [77], a general population measure of depressive symptomology, will be used to assess recent mental health. Recent substance use will be assessed using the Alcohol Use Disorders Identification Test-Alcohol Consumption Questions [82], a brief screening tool for identifying individuals with an active alcohol use disorder, and the Drug Abuse Screening Test [83].

Discussion

Study Importance

This study is designed to address gaps in knowledge about the mechanisms of HIV comorbidities among PLWHIV from health disparity populations, specifically black and Latina TWLHIV. Our transdisciplinary, multiracial team of cisgender and transgender researchers are committed to authentic, empowering engagement with transgender communities such that research is conducted with, not on, communities [107]. Leveraging the infrastructure of an existing National Institutes of Health (NIH)-funded cohort study, LITE Plus will advance scientific knowledge by applying tools from the emerging field of psychoneuroimmunology to understand how intersectional stigma and discrimination become embodied as heath inequities [108]. LITE Plus will also clarify the role of sex hormones in stress physiology by examining how GAHT may both diminish psychological distress via gender affirmation and potentially increase CVD risk because of physiologic effects of estrogen.

Despite the heavy impact of HIV on TW and the confluence of comorbidity risk factors, remarkably little research has focused on TWLHIV [109]. This study will be among the first to address

HIV comorbid conditions in this population and one of the few to prospectively assess the role of chronic stress biomarkers among PLWHIV [54,110]. In so doing, this study will advance both stigma and intersectionality research and move chronic stress physiology research toward a more nuanced understanding of sex and gender.

In filling the gap in research on GAHT among TWLHIV, this study will provide data to inform the likely complex relationship between the benefits and risks associated with estrogen in a population with an elevated risk of poor mental health and CVD related to HIV disease. Elucidating the role of sex hormones in the stress-health relationship may have implications for research with broader populations, including cisgender women. Finally, most existing studies of chronic stress, AL, and health have used cross-sectional designs [62]. Following study participants over the course of 24 months with both psychosocial and physiologic measures will strengthen the ability to make causal inferences about the nature of stress-health relationships in PLWHIV and the impact of coping and social support on health outcomes for TWLHIV, providing key data for future interventions.

Limitations, Challenges, and Solutions

Sample Size and Attrition

Budget caps limit the sample size and, therefore, statistical power, especially if attrition is significant. Retaining participants is a challenge for all longitudinal studies, and differential loss to follow-up is a common threat to internal validity in prospective research. We have developed monthly enrollment targets and extensive protocols to support retention. If we cannot reach our target sample size at current sites, the Johns Hopkins Center for Transgender Health will support additional recruitment, and we will seek supplemental funding to include additional sites from the LITE study which has locations in Atlanta, Miami, and New York as well as in Boston and the Baltimore/DC area. We are also cognizant of potential recruitment fatigue on the part of potential TW participants at these health center sites, who have likely participated in other research studies, particularly given the relatively small numbers of TW in the patient population at each site [111].

Missing Data

Home salivary collection (as well as other data collection methods) are vulnerable to missing data, despite intensive tracking protocols. Missing data reduce power and may produce biased estimates. In addition to traditional statistical approaches to handling missing data [112], we will explore growth curve modeling as an emerging method for addressing missing data in salivary cortisol research [113].

Limited Observation Period

Cumulative stress exacts a toll over time; thus 24 months may not be enough to see a change. The existing literature from the MacArthur Aging study found significant associations between AL and health over approximately 2 years, providing support for the feasibility of our 24-month study design [101].

Cardiovascular Disease Risk Assessment

Cardiac events are likely to be rare over 24 months, and budget limitations do not allow for the use of more expensive assessments such as estimation of coronary artery calcium scores. However, the CVD risk measure used for this study is common in clinical practice and is likely to have real-world applicability [114]. In addition, the absence of an existing CVD risk measure validated for use among TWLHIV is a further limitation of this study. However, the adaptation of the PCE based on GAHT exposure in this longitudinal study will make an important contribution to our understanding of CVD risk for TWLHIV. We will also conduct sensitivity analyses assessing the impact of calibrating the PCE based on sex assigned at birth, gender identity, and GAHT exposure over the study period.

Nonprescribed Hormone Use

Nonprescription GAHT use is common and may be underreported. In addition to collecting self-report and medical record data, we will conduct laboratory tests for estradiol, testosterone, and progesterone. TWLHIV taking GAHT may differ from TWLHIV not taking GAHT in ways that confound the relationship between GAHT and CVD, for example, they may have stopped GAHT because of a diagnosis of CVD. Therefore, our assessment of self-reported GAHT exposure, along with laboratory tests for sex hormones and medical record extraction of CVD-related diagnoses and medications, will allow for pragmatic assessments that are applicable to clinical practice.

Unmeasured Confounders

The proposed study involves complex, overlapping systems with multiple interacting levels and domains. In such a system, it is impossible to know and control for all potential confounders. Although unmeasured confounders may threaten external validity, the findings that result from studies that address real-world complexity (eg, by enrolling people with multimorbidity) may provide findings that are most actionable in the real world of clinical care and social policy.

Acknowledgments

The authors would like to express their gratitude to the TWLHIV who are taking part in this study. This study would not be possible without their participation. This study has been funded by the National Institute on Minority Health and Health Disparities (R01MD013498) and supported by the Centers for AIDS Research at Johns Hopkins University (P30AI094189) and Harvard University (P30AI060354). The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the NIH. The authors would like to thank the LITE study investigators, staff, and CAB members.

Conflicts of Interest

None declared.

Multimedia Appendix 1
National Institute on Minority Health and Health Disparities funded grant reviews.
[PDF File (Adobe PDF File), 162 KB - resprot_v9i4e17076_app1.pdf]

References

1. National Institute of Minority Health and Health Disparities. 2016 Oct 6. Director's Message: Sexual and Gender Minorities Formally Designated as a Health Disparity Population for Research Purposes URL: https://www.nimhd.nih.gov/about/directors-corner/message.html [accessed 2019-06-21]
2. Becasen JS, Denard CL, Mullins MM, Higa DH, Sipe TA. Estimating the prevalence of HIV and sexual behaviors among the US transgender population: a systematic review and meta-analysis, 2006-2017. Am J Public Health 2018 Nov 29:e1-e8. [doi: 10.2105/AJPH.2018.304727] [Medline: 30496000]
3. Baral SD, Poteat T, Strömdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. Lancet Infect Dis 2013 Mar;13(3):214-222. [doi: 10.1016/S1473-3099(12)70315-8] [Medline: 23260128]
4. Poteat T, Scheim A, Xavier J, Reisner S, Baral S. Global epidemiology of HIV infection and related syndemics affecting transgender people. J Acquir Immune Defic Syndr 2016 Aug 15;72(Suppl 3):S210-S219 [FREE Full text] [doi: 10.1097/QAI.0000000000001087] [Medline: 27429185]
5. Herbst JH, Jacobs ED, Finlayson TJ, McKleroy VS, Neumann MS, Crepaz N, HIV/AIDS Prevention Research Synthesis Team. Estimating HIV prevalence and risk behaviors of transgender persons in the United States: a systematic review. AIDS Behav 2008 Jan;12(1):1-17. [doi: 10.1007/s10461-007-9299-3] [Medline: 17694429]
6. Health Resources & Services Administration. Ryan White HIV/AIDS Program Annual Client-Level Data Report 2015 URL: http://hab.hrsa.gov/data/data-reports [accessed 2017-03-24]
7. Mizuno Y, Frazier EL, Huang P, Skarbinski J. Characteristics of transgender women living with HIV receiving medical care in the United States. LGBT Health 2015 Sep;2(3):228-234 [FREE Full text] [doi: 10.1089/lgbt.2014.0099] [Medline: 26788671]
8. Reisner SL, Poteat T, Keatley J, Cabral M, Mothopeng T, Dunham E, et al. Global health burden and needs of transgender populations: a review. Lancet 2016 Jul 23;388(10042):412-436. [doi: 10.1016/S0140-6736(16)00684-X] [Medline: 27323919]
9. Poteat T, Reisner SL, Radix A. HIV epidemics among transgender women. Curr Opin HIV AIDS 2014 Mar;9(2):168-173 [FREE Full text] [doi: 10.1097/COH.0000000000000030] [Medline: 24322537]
10. Brown GR, Jones KT. Mental health and medical health disparities in 5135 transgender veterans receiving healthcare in the Veterans Health Administration: a case-control study. LGBT Health 2016 Apr;3(2):122-131. [doi: 10.1089/lgbt.2015.0058] [Medline: 26674598]
11. Radix A, Sevelius J, Deutsch MB. Transgender women, hormonal therapy and HIV treatment: a comprehensive review of the literature and recommendations for best practices. J Int AIDS Soc 2016;19(3 Suppl 2):20810 [FREE Full text] [doi: 10.7448/IAS.19.3.20810] [Medline: 27431475]
12. Rai S, Mahapatra B, Sircar S, Raj PY, Venkatesh S, Shaukat M, et al. Adherence to antiretroviral therapy and its effect on survival of HIV-infected individuals in Jharkhand, India. PLoS One 2013;8(6):e66860 [FREE Full text] [doi: 10.1371/journal.pone.0066860] [Medline: 23825577]
13. James SE, Herman JL, Rankin S, Keisling M, Mottet L, Anafi M. National Center for Transgender Equality. 2016. The Report of the 2015 US Transgender Survey URL: https://transequality.org/sites/default/files/docs/usts/USTS-Full-Report-Dec17.pdf [accessed 2018-01-28]
14. Nuttbrock L, Bockting W, Rosenblum A, Hwahng S, Mason M, Macri M, et al. Gender abuse, depressive symptoms, and HIV and other sexually transmitted infections among male-to-female transgender persons: a three-year prospective study. Am J Public Health 2013 Feb;103(2):300-307. [doi: 10.2105/AJPH.2011.300568] [Medline: 22698023]
15. Nemoto T, Bödeker B, Iwamoto M. Social support, exposure to violence and transphobia, and correlates of depression among male-to-female transgender women with a history of sex work. Am J Public Health 2011 Oct;101(10):1980-1988. [doi: 10.2105/AJPH.2010.197285] [Medline: 21493940]
16. Bockting WO, Miner MH, Swinburne Romine RE, Hamilton A, Coleman E. Stigma, mental health, and resilience in an online sample of the US transgender population. Am J Public Health 2013 May;103(5):943-951. [doi: 10.2105/AJPH.2013.301241] [Medline: 23488522]
17. Perez-Brumer A, Hatzenbuehler ML, Oldenburg CE, Bockting W. Individual- and structural-level risk factors for suicide attempts among transgender adults. Behav Med 2015;41(3):164-171 [FREE Full text] [doi: 10.1080/08964289.2015.1028322] [Medline: 26287284]
18. Nanni MG, Caruso R, Mitchell AJ, Meggiolaro E, Grassi L. Depression in HIV infected patients: a review. Curr Psychiatry Rep 2015 Jan;17(1):530. [doi: 10.1007/s11920-014-0530-4] [Medline: 25413636]

19. Bradley SM, Rumsfeld JS. Depression and cardiovascular disease. Trends Cardiovasc Med 2015 Oct;25(7):614-622. [doi: 10.1016/j.tcm.2015.02.002] [Medline: 25850976]
20. Mizuno Y, Beer L, Huang P, Frazier EL. Factors associated with antiretroviral therapy adherence among transgender women receiving HIV medical care in the United States. LGBT Health 2017 Jun;4(3):181-187 [FREE Full text] [doi: 10.1089/lgbt.2017.0003] [Medline: 28498011]
21. Asscheman H, Giltay E, Megens J, de Ronde WP, van Trotsenburg MA, Gooren L. A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. Eur J Endocrinol 2011 Apr;164(4):635-642. [doi: 10.1530/EJE-10-1038] [Medline: 21266549]
22. Wierckx K, Elaut E, Declercq E, Heylens G, de Cuypere G, Taes Y, et al. Prevalence of cardiovascular disease and cancer during cross-sex hormone therapy in a large cohort of trans persons: a case-control study. Eur J Endocrinol 2013 Oct;169(4):471-478. [doi: 10.1530/EJE-13-0493] [Medline: 23904280]
23. So-Armah K, Freiberg MS. Cardiovascular disease risk in an aging HIV population: not just a question of biology. Curr Opin HIV AIDS 2014 Jul;9(4):346-354 [FREE Full text] [doi: 10.1097/COH.0000000000000065] [Medline: 24824885]
24. Nemeth C, Bekhbat M, Neigh G. Neural effects of inflammation, cardiovascular disease, and HIV: parallel, perpendicular, or progressive? Neuroscience 2015 Aug 27;302:165-173 [FREE Full text] [doi: 10.1016/j.neuroscience.2014.09.016] [Medline: 25239371]
25. Gamarel KE, Mereish EH, Manning D, Iwamoto M, Operario D, Nemoto T. Minority stress, smoking patterns, and cessation attempts: findings from a community-sample of transgender women in the San Francisco bay area. Nicotine Tob Res 2016 Mar;18(3):306-313 [FREE Full text] [doi: 10.1093/ntr/ntv066] [Medline: 25782458]
26. Maraka S, Ospina NS, Rodriguez-Gutierrez R, Davidge-Pitts CJ, Nippoldt TB, Prokop LJ, et al. Sex steroids and cardiovascular outcomes in transgender individuals: a systematic review and meta-analysis. J Clin Endocrinol Metab 2017 Nov 1;102(11):3914-3923. [doi: 10.1210/jc.2017-01643] [Medline: 28945852]
27. Sanchez NF, Sanchez JP, Danoff A. Health care utilization, barriers to care, and hormone usage among male-to-female transgender persons in New York City. Am J Public Health 2009 Apr;99(4):713-719. [doi: 10.2105/AJPH.2007.132035] [Medline: 19150911]
28. de Haan G, Santos G, Arayasirikul S, Raymond HF. Non-prescribed hormone use and barriers to care for transgender women in San Francisco. LGBT Health 2015 Dec;2(4):313-323. [doi: 10.1089/lgbt.2014.0128] [Medline: 26788772]
29. Beckwith N, Reisner SL, Zaslow S, Mayer KH, Keuroghlian AS. Factors associated with gender-affirming surgery and age of hormone therapy initiation among transgender adults. Transgend Health 2017;2(1):156-164 [FREE Full text] [doi: 10.1089/trgh.2017.0028] [Medline: 29159310]
30. Bauer GR, Scheim AI, Pyne J, Travers R, Hammond R. Intervenable factors associated with suicide risk in transgender persons: a respondent driven sampling study in Ontario, Canada. BMC Public Health 2015 Jun 2;15:525 [FREE Full text] [doi: 10.1186/s12889-015-1867-2] [Medline: 26032733]
31. Wilson EC, Chen Y, Arayasirikul S, Wenzel C, Raymond HF. Connecting the dots: examining transgender women's utilization of transition-related medical care and associations with mental health, substance use, and HIV. J Urban Health 2015 Feb;92(1):182-192 [FREE Full text] [doi: 10.1007/s11524-014-9921-4] [Medline: 25476958]
32. Sevelius JM, Patouhas E, Keatley JG, Johnson MO. Barriers and facilitators to engagement and retention in care among transgender women living with human immunodeficiency virus. Ann Behav Med 2014 Feb;47(1):5-16 [FREE Full text] [doi: 10.1007/s12160-013-9565-8] [Medline: 24317955]
33. Sevelius JM, Carrico A, Johnson MO. Antiretroviral therapy adherence among transgender women living with HIV. J Assoc Nurses AIDS Care 2010;21(3):256-264 [FREE Full text] [doi: 10.1016/j.jana.2010.01.005] [Medline: 20347342]
34. Sevelius JM, Saberi P, Johnson MO. Correlates of antiretroviral adherence and viral load among transgender women living with HIV. AIDS Care 2014;26(8):976-982 [FREE Full text] [doi: 10.1080/09540121.2014.896451] [Medline: 24646419]
35. Deutsch M, Chakravarty D, Rebchook G, Shade S, Sevelius J, Maiorana A. Associations between self-reported hormone use patterns and indicators of HIV care among transgender women of color in four U.S. cities (abstract ID 1886). In: Associations between self-reported hormone use patterns and indicators of HIV care among transgender women of color in four U.S. cities (abstract ID 1886). Atlanta, GA: Centers for Disease Control and Prevention (CDC); 2015 Presented at: 2015 National HIV Prevention Conference; December 6-9, 2015; Atlanta, GA p. 1886.
36. Streed CG, Harfouch O, Marvel F, Blumenthal RS, Martin SS, Mukherjee M. Cardiovascular disease among transgender adults receiving hormone therapy: a narrative review. Ann Intern Med 2017 Aug 15;167(4):256-267. [doi: 10.7326/M17-0577] [Medline: 28738421]
37. Klein LC, Corwin EJ. Seeing the unexpected: how sex differences in stress responses may provide a new perspective on the manifestation of psychiatric disorders. Curr Psychiatry Rep 2002 Dec;4(6):441-448. [doi: 10.1007/s11920-002-0072-z] [Medline: 12441024]
38. Kajantie E, Phillips DI. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. Psychoneuroendocrinology 2006 Feb;31(2):151-178. [doi: 10.1016/j.psyneuen.2005.07.002] [Medline: 16139959]
39. Williams GP. The role of oestrogen in the pathogenesis of obesity, type 2 diabetes, breast cancer and prostate disease. Eur J Cancer Prev 2010 Jul;19(4):256-271. [doi: 10.1097/cej.0b013e328338f7d2] [Medline: 20535861]

40. Martel MM. Sexual selection and sex differences in the prevalence of childhood externalizing and adolescent internalizing disorders. Psychol Bull 2013 Nov;139(6):1221-1259. [doi: 10.1037/a0032247] [Medline: 23627633]
41. McEwen BS, Stellar E. Stress and the individual. Mechanisms leading to disease. Arch Intern Med 1993 Sep 27;153(18):2093-2101. [doi: 10.1001/archinte.1993.00410180039004] [Medline: 8379800]
42. McEwen BS. Stress: homeostasis, rheostasis, allostasis and allostatic load. In: Squire LR, Bloom FE, Spitzer NC, Gage F, Albright T, editors. Encyclopedia of Neuroscience. Oxford: Academic Press; 2009:557-561.
43. Juster R, Russell JJ, Almeida D, Picard M. Allostatic load and comorbidities: a mitochondrial, epigenetic, and evolutionary perspective. Dev Psychopathol 2016 Nov;28(4pt1):1117-1146. [doi: 10.1017/S0954579416000730] [Medline: 27739386]
44. Sterling P, Eyer J. Allostasis: a new paradigm to explain arousal pathology. In: Fisher S, editor. Handbook of Life Stress, Cognition and Health. New York: John Wiley & Sons; 1988:629-649.
45. McEwen BS. Allostasis and allostatic load: implications for neuropsychopharmacology. Neuropsychopharmacology 2000 Feb;22(2):108-124 [FREE Full text] [doi: 10.1016/S0893-133X(99)00129-3] [Medline: 10649824]
46. Lucas T, Wegner R, Pierce J, Lumley MA, Laurent HK, Granger DA. Perceived discrimination, racial identity, and multisystem stress response to social evaluative threat among African American men and women. Psychosom Med 2017 Apr;79(3):293-305 [FREE Full text] [doi: 10.1097/PSY.0000000000000406] [Medline: 27806018]
47. Jackson JS, Knight KM, Rafferty JA. Race and unhealthy behaviors: chronic stress, the HPA axis, and physical and mental health disparities over the life course. Am J Public Health 2010 May;100(5):933-939 [FREE Full text] [doi: 10.2105/AJPH.2008.143446] [Medline: 19846689]
48. Busse D, Yim IS, Campos B. Social context matters: ethnicity, discrimination and stress reactivity. Psychoneuroendocrinology 2017 Sep;83:187-193. [doi: 10.1016/j.psyneuen.2017.05.025] [Medline: 28646798]
49. Gallo LC, Roesch SC, Fortmann AL, Carnethon MR, Penedo FJ, Perreira K, et al. Associations of chronic stress burden, perceived stress, and traumatic stress with cardiovascular disease prevalence and risk factors in the Hispanic Community Health Study/Study of Latinos Sociocultural Ancillary Study. Psychosom Med 2014;76(6):468-475 [FREE Full text] [doi: 10.1097/PSY.0000000000000069] [Medline: 24979579]
50. Hatzenbuehler ML, Slopen N, McLaughlin KA, McLaughlin KA. Stressful life events, sexual orientation, and cardiometabolic risk among young adults in the United States. Health Psychol 2014 Oct;33(10):1185-1194 [FREE Full text] [doi: 10.1037/hea0000126] [Medline: 25133830]
51. Juster R, Smith NG, Ouellet E, Sindi S, Lupien SJ. Sexual orientation and disclosure in relation to psychiatric symptoms, diurnal cortisol, and allostatic load. Psychosom Med 2013 Feb;75(2):103-116. [doi: 10.1097/PSY.0b013e3182826881] [Medline: 23362500]
52. Dubois LZ. Associations between transition-specific stress experience, nocturnal decline in ambulatory blood pressure, and C-reactive protein levels among transgender men. Am J Hum Biol 2012;24(1):52-61 [FREE Full text] [doi: 10.1002/ajhb.22203] [Medline: 22120883]
53. DuBois LZ, Powers S, Everett BG, Juster R. Stigma and diurnal cortisol among transitioning transgender men. Psychoneuroendocrinology 2017 Aug;82:59-66. [doi: 10.1016/j.psyneuen.2017.05.008] [Medline: 28511045]
54. Glover DA, Garcia-Aracena EF, Lester P, Rice E, Rothram-Borus MJ. Stress biomarkers as outcomes for HIV+ prevention: participation, feasibility and findings among HIV+ Latina and African American mothers. AIDS Behav 2010 Apr;14(2):339-350 [FREE Full text] [doi: 10.1007/s10461-009-9549-7] [Medline: 19350378]
55. Bowleg L. The problem with the phrase women and minorities: intersectionality-an important theoretical framework for public health. Am J Public Health 2012 Jul;102(7):1267-1273. [doi: 10.2105/AJPH.2012.300750] [Medline: 22594719]
56. Bauer GR. Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. Soc Sci Med 2014 Jun;110:10-17 [FREE Full text] [doi: 10.1016/j.socscimed.2014.03.022] [Medline: 24704889]
57. Sangaramoorthy T, Jamison A, Dyer T. Intersectional stigma among midlife and older Black women living with HIV. Cult Health Sex 2017 Dec;19(12):1329-1343 [FREE Full text] [doi: 10.1080/13691058.2017.1312530] [Medline: 28418279]
58. Testa RJ, Habarth J, Peta J, Balsam K, Bockting W. Development of the gender minority stress and resilience measure. Psychol Sex Orientat Gend Divers 2015;2(1):65-77. [doi: 10.1037/sgd0000081]
59. Hendricks ML, Testa RJ. A conceptual framework for clinical work with transgender and gender nonconforming clients: an adaptation of the Minority Stress Model. Prof Psychol Res Pract 2012;43(5):460-467. [doi: 10.1037/a0029597]
60. Stewart JA. The detrimental effects of allostasis: allostatic load as a measure of cumulative stress. J Physiol Anthropol 2006 Jan;25(1):133-145 [FREE Full text] [doi: 10.2114/jpa2.25.133] [Medline: 16617218]
61. Carlson ED, Chamberlain RM. Allostatic load and health disparities: a theoretical orientation. Res Nurs Health 2005 Aug;28(4):306-315. [doi: 10.1002/nur.20084] [Medline: 16028266]
62. Beckie TM. A systematic review of allostatic load, health, and health disparities. Biol Res Nurs 2012 Oct;14(4):311-346. [doi: 10.1177/1099800412455688] [Medline: 23007870]
63. Alexander-Floyd NG. Disappearing acts: reclaiming intersectionality in the social sciences in a post-Black feminist era. Feminist Formations 2012;24(1):1-25. [doi: 10.1353/ff.2012.0003]
64. Wirtz AL, Poteat T, Radix A, Althoff KN, Cannon CM, Wawrzyniak AJ, American Cohort To Study HIV Acquisition Among Transgender Women (LITE). American cohort to study HIV acquisition among transgender women in high-risk

areas (The LITE Study): protocol for a multisite prospective cohort study in the Eastern and Southern United States. JMIR Res Protoc 2019 Oct 3;8(10):e14704 [FREE Full text] [doi: 10.2196/14704] [Medline: 31584005]
65. Reynolds NR, Sun J, Nagaraja HN, Gifford AL, Wu AW, Chesney MA. Optimizing measurement of self-reported adherence with the ACTG Adherence Questionnaire: a cross-protocol analysis. J Acquir Immune Defic Syndr 2007 Dec 1;46(4):402-409. [doi: 10.1097/qai.0b013e318158a44f] [Medline: 18077832]
66. Taylor JM. Psychometric analysis of the Ten-Item Perceived Stress Scale. Psychol Assess 2015 Mar;27(1):90-101. [doi: 10.1037/a0038100] [Medline: 25346996]
67. Sinclair VG, Wallston KA. The development and psychometric evaluation of the Brief Resilient Coping Scale. Assessment 2004 Mar;11(1):94-101. [doi: 10.1177/1073191103258144] [Medline: 14994958]
68. Newschaffer CJ. Centers for Disease Control and Prevention. Atlanta, GA: Centers for Disease Control and Prevention; 1998. Validation of Behavioral Risk Factor Surveillance System (BRFSS) HRQOL Measures in a Statewide Sample URL: https://www.cdc.gov/hrqol/pdfs/validationreport.pdf [accessed 2019-06-01]
69. Arbona C, Olvera N, Rodriguez N, Hagan J, Linares A, Wiesner M. Acculturative stress among documented and undocumented Latino immigrants in the United States. Hisp J Behav Sci 2010 Aug;32(3):362-384 [FREE Full text] [doi: 10.1177/0739986310373210] [Medline: 25484488]
70. Rodriguez N, Paredes CL, Hagan J. Fear of immigration enforcement among older Latino immigrants in the United States. J Aging Health 2017 Sep;29(6):986-1014. [doi: 10.1177/0898264317710839] [Medline: 28670946]
71. Sayles JN, Hays RD, Sarkisian CA, Mahajan AP, Spritzer KL, Cunningham WE. Development and psychometric assessment of a multidimensional measure of internalized HIV stigma in a sample of HIV-positive adults. AIDS Behav 2008 Sep;12(5):748-758 [FREE Full text] [doi: 10.1007/s10461-008-9375-3] [Medline: 18389363]
72. Scheim AI, Bauer GR. The Intersectional Discrimination Index: Development and validation of measures of self-reported enacted and anticipated discrimination for intercategorical analysis. Soc Sci Med 2019 Apr;226:225-235 [FREE Full text] [doi: 10.1016/j.socscimed.2018.12.016] [Medline: 30674436]
73. Sternthal MJ, Slopen N, Williams DR. Racial disparities in health: how much does stress really matter? Du Bois Rev 2011;8(1):95-113 [FREE Full text] [doi: 10.1017/S1742058X11000087] [Medline: 29887911]
74. Sevelius J, Chakravarty D, Neilands TB, Keatley J, Shade SB, Johnson MO, HRSA SPNS Transgender Women of Color Study Group. Evidence for the model of gender affirmation: the role of gender affirmation and healthcare empowerment in viral suppression among transgender women of color living with HIV. AIDS Behav 2019 May 29. [doi: 10.1007/s10461-019-02544-2] [Medline: 31144131]
75. Safran DG, Kosinski M, Tarlov AR, Rogers WH, Taira DA, Lieberman N, et al. The Primary Care Assessment Survey: tests of data quality and measurement performance. Med Care 1998 May;36(5):728-739. [doi: 10.1097/00005650-199805000-00012] [Medline: 9596063]
76. Weathers F, Litz B, Herman D, Huska J, Keane T. ResearchGate. 1993. The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility URL: https://www.researchgate.net/publication/291448760_The_PTSD_Checklist_PCL_Reliability_validity_and_diagnostic_utility [accessed 2020-02-10]
77. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [doi: 10.1016/s0749-3797(18)30622-6] [Medline: 8037935]
78. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The cardiovascular health study: design and rationale. Ann Epidemiol 1991 Feb;1(3):263-276. [doi: 10.1016/1047-2797(91)90005-w] [Medline: 1669507]
79. National Cancer Institute: Comprehensive Cancer Information. NCI Dictionary of Cancer Terms: Pack Year URL: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/pack-year?redirect=true [accessed 2019-10-30]
80. California Health Interview Survey (CHIS). UCLA Center for Health Policy Research. Los Angeles, CA: UCLA Center for Health Policy Research; 2011 Aug 30. CHIS 2003 Adult Questionnaire URL: https://healthpolicy.ucla.edu/chis/design/Documents/_CHIS2003_adult_q.pdf [accessed 2020-02-07]
81. Hardie JH, Lucas A. Economic factors and relationship quality among young couples: Comparing cohabitation and marriage. J Marriage Fam 2010 Oct;72(5):1141-1154. [doi: 10.1111/j.1741-3737.2010.00755.x] [Medline: 21691414]
82. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
83. Skinner HA. The drug abuse screening test. Addict Behav 1982;7(4):363-371. [doi: 10.1016/0306-4603(82)90005-3] [Medline: 7183189]
84. Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The revised conflict tactics scales (CTS2): development and preliminary psychometric data. J Fam Issues 1996;17(3):283-316. [doi: 10.1177/019251396017003001]
85. Granger DA, Kivlighan KT, Fortunato C, Harmon AG, Hibel LC, Schwartz EB, et al. Integration of salivary biomarkers into developmental and behaviorally-oriented research: problems and solutions for collecting specimens. Physiol Behav 2007 Nov 23;92(4):583-590. [doi: 10.1016/j.physbeh.2007.05.004] [Medline: 17572453]

86. Peng H, Long Y, Li J, Guo Y, Wu H, Yang Y, et al. Hypothalamic-pituitary-adrenal axis functioning and dysfunctional attitude in depressed patients with and without childhood neglect. BMC Psychiatry 2014 Feb 18;14:45 [FREE Full text] [doi: 10.1186/1471-244X-14-45] [Medline: 24548345]
87. Pico-Alfonso MA, Garcia-Linares MI, Celda-Navarro N, Herbert J, Martinez M. Changes in cortisol and dehydroepiandrosterone in women victims of physical and psychological intimate partner violence. Biol Psychiatry 2004 Aug 15;56(4):233-240. [doi: 10.1016/j.biopsych.2004.06.001] [Medline: 15312810]
88. Inder WJ, Dimeski G, Russell A. Measurement of salivary cortisol in 2012 - laboratory techniques and clinical indications. Clin Endocrinol (Oxf) 2012 Nov;77(5):645-651. [doi: 10.1111/j.1365-2265.2012.04508.x] [Medline: 22812714]
89. Lee DY, Kim E, Choi MH. Technical and clinical aspects of cortisol as a biochemical marker of chronic stress. BMB Rep 2015 Apr;48(4):209-216 [FREE Full text] [doi: 10.5483/bmbrep.2015.48.4.275] [Medline: 25560699]
90. Granger DA, Johnson SB, Szanton SL, Out D, Schumann LL. Incorporating salivary biomarkers into nursing research: an overview and review of best practices. Biol Res Nurs 2012 Oct;14(4):347-356 [FREE Full text] [doi: 10.1177/1099800412443892] [Medline: 22593229]
91. Hodgson NA, Granger DA. Collecting saliva and measuring salivary cortisol and alpha-amylase in frail community residing older adults via family caregivers. J Vis Exp 2013 Dec 18(82):e50815 [FREE Full text] [doi: 10.3791/50815] [Medline: 24378361]
92. Goldenberg T, Finneran C, Andes KL, Stephenson R. Using participant-empowered visual relationship timelines in a qualitative study of sexual behaviour. Glob Public Health 2016;11(5-6):699-718. [doi: 10.1080/17441692.2016.1170869] [Medline: 27092985]
93. Wirtz AL, Cooney EE, Chaudhry A, Reisner SL, American Cohort To Study HIV Acquisition Among Transgender Women. Computer-mediated communication to facilitate synchronous online focus group discussions: feasibility study for qualitative HIV research among transgender women across the United States. J Med Internet Res 2019 Mar 29;21(3):e12569 [FREE Full text] [doi: 10.2196/12569] [Medline: 30924782]
94. Belli RF, James S, van Hoewyk J, Alcser KH. The implementation of a computerized event history calendar questionnaire for research in life course epidemiology. In: Stafford FP, Belli RF, Alwin DF, editors. Calendar and Time Diary Methods in Life Course Research. Thousand Oaks, CA: SAGE Publications, Inc; 2009:224-238.
95. Axinn WG, Pearce LD, Ghimire D. Innovations in life history calendar applications. Soc Sci Res 1999;28(3):243-264. [doi: 10.1006/ssre.1998.0641]
96. Althoff KN, Rebeiro P, Brooks JT, Buchacz K, Gebo K, Martin J, North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD). Disparities in the quality of HIV care when using US Department of Health and Human Services indicators. Clin Infect Dis 2014 Apr;58(8):1185-1189 [FREE Full text] [doi: 10.1093/cid/ciu044] [Medline: 24463281]
97. Gatsonis C, Sampson AR. Multiple correlation: exact power and sample size calculations. Psychol Bull 1989 Nov;106(3):516-524. [doi: 10.1037/0033-2909.106.3.516] [Medline: 2813654]
98. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
99. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013 Jun;18(2):137-150 [FREE Full text] [doi: 10.1037/a0031034] [Medline: 23379553]
100. Elder Jr GH, Johnson MK, Crosnoe R. The emergence and development of life course theory. In: Handbook of the Life Course. Boston, MA: Springer; 2003:3-19.
101. Seeman TE, Singer BH, Rowe JW, Horwitz RI, McEwen BS. Price of adaptation--allostatic load and its health consequences. MacArthur studies of successful aging. Arch Intern Med 1997 Oct 27;157(19):2259-2268. [Medline: 9343003]
102. Juster RP, McEwen BS, Lupien SJ. Allostatic load biomarkers of chronic stress and impact on health and cognition. Neurosci Biobehav Rev 2010 Sep;35(1):2-16. [doi: 10.1016/j.neubiorev.2009.10.002] [Medline: 19822172]
103. Seeman TE, McEwen BS, Rowe JW, Singer BH. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. Proc Natl Acad Sci USA 2001 Apr 10;98(8):4770-4775 [FREE Full text] [doi: 10.1073/pnas.081072698] [Medline: 11287659]
104. Juster RP, Bizik G, Picard M, Arsenault-Lapierre G, Sindi S, Trepanier L, et al. A transdisciplinary perspective of chronic stress in relation to psychopathology throughout life span development. Dev Psychopathol 2011 Aug;23(3):725-776. [doi: 10.1017/S0954579411000289] [Medline: 21756430]
105. American College of Cardiology. ASCVD Risk Estimator Plus URL: http://tools.acc.org/ASCVD-Risk-Estimator-Plus/#!/calculate/estimate/ [accessed 2019-07-30]
106. Reisner SL, Biello KB, Hughto JM, Kuhns L, Mayer KH, Garofalo R, et al. Psychiatric diagnoses and comorbidities in a diverse, multicity cohort of young transgender women: baseline findings from project lifeskills. JAMA Pediatr 2016 May 1;170(5):481-486 [FREE Full text] [doi: 10.1001/jamapediatrics.2016.0067] [Medline: 26999485]
107. Scheim A, Appenroth M, Beckham S, Goldstein Z, Grinspan M, Keatley J, et al. Transgender HIV research: nothing about us without us. The Lancet HIV 2019 Sep;6(9):e566-e567. [doi: 10.1016/S2352-3018(19)30269-3]

108. Krieger N. Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination. Int J Health Serv 1999;29(2):295-352. [doi: 10.2190/M11W-VWXE-KQM9-G97Q] [Medline: 10379455]
109. Poteat TC, Hanna DB, Althoff KN. Short communication: feasibility and acceptability of developing a multisite clinical cohort of transgender people with HIV infection. AIDS Res Hum Retroviruses 2015 Sep;31(9):870-872 [FREE Full text] [doi: 10.1089/aid.2015.0055] [Medline: 26126154]
110. Glover DA, Williams JK, Kisler KA. Using novel methods to examine stress among HIV-positive African American men who have sex with men and women. J Behav Med 2013 Jun;36(3):283-294 [FREE Full text] [doi: 10.1007/s10865-012-9421-5] [Medline: 22538773]
111. Poteat T, Wirtz A, Malik M, Cooney E, Cannon C, Hardy WD, et al. A gap between willingness and uptake: Findings from mixed methods research on HIV prevention among Black and Latina transgender women. J Acquir Immune Defic Syndr 2019 Oct 1;82(2):131-140. [doi: 10.1097/QAI.0000000000002112] [Medline: 31180995]
112. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol 2013 May;64(5):402-406 [FREE Full text] [doi: 10.4097/kjae.2013.64.5.402] [Medline: 23741561]
113. Hogue CM, Pornprasertmanit S, Fry MD, Rhemtulla M, Little TD. Planned missing data designs for spline growth models in salivary cortisol research. Meas Phys Educ Exerc Sci 2013;17(4):310-325. [doi: 10.1080/1091367x.2013.831766]
114. Topel ML, Shen J, Morris AA, Al Mheid I, Sher S, Dunbar SB, et al. Comparisons of the Framingham and pooled cohort equation risk scores for detecting subclinical vascular disease in blacks versus whites. Am J Cardiol 2018 Mar 1;121(5):564-569 [FREE Full text] [doi: 10.1016/j.amjcard.2017.11.031] [Medline: 29361288]

Abbreviations

AL: allostatic load
ALI: allostatic load index
ART: antiretroviral therapy
BLTW: black and Latina transgender women
CAB: community advisory board
CRP: C-reactive protein
CTMS: Clinical Trials Management System
CVD: cardiovascular disease
DHEAS: dehydroepiandrosterone-sulphate
GAHT: gender-affirming hormone therapy
GMSR: Gender Minority Stress and Resilience Model
Hb: hemoglobin
HDL: high-density lipoprotein
HIPAA: Health Insurance Portability and Accountability Act
IL-6: interleukin-6
LHC: life history calendar
LITE: American Cohort to Study HIV Acquisition among Transgender Women in High Risk Areas
NIH: National Institutes of Health
PCE: pooled cohort equations
PLWHIV: people living with HIV
TNF-alpha: tumor necrosis factor-alpha
TW: transgender women
TWLHIV: transgender women living with HIV

Edited by G Eysenbach; submitted 15.11.19; peer-reviewed by A Eaton, E Olansky; comments to author 11.01.20; revised version received 23.01.20; accepted 24.01.20; published 13.04.20.

Please cite as:
Rich AJ, Williams J, Malik M, Wirtz A, Reisner S, DuBois LZ, Juster RP, Lesko CR, Davis N, Althoff KN, Cannon C, Mayer K, Elliott A, Poteat T
Biopsychosocial Mechanisms Linking Gender Minority Stress to HIV Comorbidities Among Black and Latina Transgender Women (LITE Plus): Protocol for a Mixed Methods Longitudinal Study
JMIR Res Protoc 2020;9(4):e17076
URL: http://www.researchprotocols.org/2020/4/e17076/
doi:10.2196/17076
PMID:32281542

Protocol

Nursing in the Age of Artificial Intelligence: Protocol for a Scoping Review

Christine Buchanan[1], BNSc, MN; M Lyndsay Howitt[1], BScN, MPH; Rita Wilson[1], BScN, MN, MEd; Richard G Booth[2], BScN, MScN, PhD; Tracie Risling[3], BA, BSN, MN, PhD; Megan Bamford[1], BScN, MScN

[1]Registered Nurses' Association of Ontario, Toronto, ON, Canada

[2]Arthur Labatt Family School of Nursing, Western University, London, ON, Canada

[3]College of Nursing, University of Saskatchewan, Saskatoon, SK, Canada

Corresponding Author:
Christine Buchanan, BNSc, MN
Registered Nurses' Association of Ontario
158 Pearl Street
Toronto, ON, M5H 1L3
Canada
Phone: 1 800 268 7199 ext 281
Email: cbuchanan@rnao.ca

Abstract

Background: It is predicted that digital health technologies that incorporate artificial intelligence will transform health care delivery in the next decade. Little research has explored how emerging trends in artificial intelligence–driven digital health technologies may influence the relationship between nurses and patients.

Objective: The purpose of this scoping review is to summarize the findings from 4 research questions regarding emerging trends in artificial intelligence–driven digital health technologies and their influence on nursing practice across the 5 domains outlined by the Canadian Nurses Association framework: administration, clinical care, education, policy, and research. Specifically, this scoping review will examine how emerging trends will transform the roles and functions of nurses over the next 10 years and beyond.

Methods: Using an established scoping review methodology, MEDLINE, Cumulative Index to Nursing and Allied Health Literature, Embase, PsycINFO, Cochrane Database of Systematic Reviews, Cochrane Central, Education Resources Information Centre, Scopus, Web of Science, and Proquest databases were searched. In addition to the electronic database searches, a targeted website search will be performed to access relevant grey literature. Abstracts and full-text studies will be independently screened by 2 reviewers using prespecified inclusion and exclusion criteria. Included literature will focus on nursing and digital health technologies that incorporate artificial intelligence. Data will be charted using a structured form and narratively summarized.

Results: Electronic database searches have retrieved 10,318 results. The scoping review and subsequent briefing paper will be completed by the fall of 2020.

Conclusions: A symposium will be held to share insights gained from this scoping review with key thought leaders and a cross section of stakeholders from administration, clinical care, education, policy, and research as well as patient advocates. The symposium will provide a forum to explore opportunities for action to advance the future of nursing in a technological world and, more specifically, nurses' delivery of compassionate care in the age of artificial intelligence. Results from the symposium will be summarized in the form of a briefing paper and widely disseminated to relevant stakeholders.

International Registered Report Identifier (IRRID): DERR1-10.2196/17490

(JMIR Res Protoc 2020;9(4):e17490) doi:10.2196/17490

KEYWORDS

nursing; artificial intelligence; machine learning; robotics; compassionate care; scoping review

Introduction

Artificial Intelligence

Artificial intelligence (AI) has been defined as "the theory and development of computer systems [which are] able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages" [1]. Machine learning and deep learning are subsets of AI and have become popular neologisms used to describe various, more specific algorithmic methodologies and techniques used to process information in ways that can imitate human-like decision making [2]. Within health care, the use of technology possessing AI has become increasingly popularized due to its capacity to sort through, analyze, and find patterns among large amounts of research evidence and patient data, ultimately discovering new meaning [3]. Current examples of where AI has been integrated within health care include clinical decision support systems [4,5], virtual nurses [6], and social robots with natural language processing abilities [7]. Furthermore, researchers are currently exploring the use of deep learning for diagnostic purposes and prediction of future clinical events [8], and machine learning is being used to understand past patient experiences, while constantly changing and updating as new data become available [3]. Given the enormity of the financial investments in these technologies made to date, it is predicted that health care services and delivery will be transformed in significant ways in the coming decade; authors have suggested that global spending on digital health technologies that incorporate AI will exceed $36 billion by 2025 [9]. In addition, the paradigm shift in preparing nurses for the digital future is already being discussed; one recent publication suggests that AI-driven technologies need to be considered as a new means of addressing health care challenges in the 21st century and that the health care workforce needs to be prepared for these changes accordingly [10]. Although AI is still a nascent topic, emerging literature has suggested that digital health technologies that use elements of AI will begin to impact the daily aspects of people's lives in the not-too-distant future [11,12].

Background

AI-driven digital health technologies that have decision-making capacities independent of humans are currently being used in numerous health care organizations. For example, early warning systems and clinical decision support systems that utilize machine learning principles are being used to aid nursing workflow and provide more personalized patient care in hospital settings [11]. Virtual nurses are another example of a digital health technology that incorporates AI. Virtual nursing avatars are accessible through a patient's computer or smartphone device and enable health care organizations to collect patient information, provide discharge instructions, coach patients, and assess patient health status from a remote location [6]. In addition, there is a growing body of literature exploring the use of intelligent assistive technological devices to support older adults that live alone with various activities of daily living [13]. Such devices have the capacity to sense and respond to consumer needs, and their intelligent nature allows them to operate autonomously within a network of related devices [13]. Not only are these technologies relevant to older adults in the home setting but they can also be translated for use in other health care contexts, such as long-term care residences and rehabilitation centers [7]. Finally, in Canada, Humber River Hospital has incorporated a fully digital infrastructure that includes the use of AI-driven robots to support children and their families as part of the Child Life Program [14]. These examples demonstrate the capabilities of AI and the ways in which this technology is already starting to transform health care systems.

While AI is frequently used in health care to assist with data analytics and clinical decision making [11], the potential for AI-driven digital health technologies to influence the relationship between nurses and their patients must not be understated. As examined by Idhe's philosophy of technology [15], there is invariably a relationship between human beings, technology, and the surrounding world; whether technology is working in the background of our daily lives or through embodied relationships, there are numerous human-technology relationships that can be explored further [15]. Thus, this scoping review is intended to further one's understanding of how AI-driven health technologies influence the nurse-patient relationship and nurse work in general.

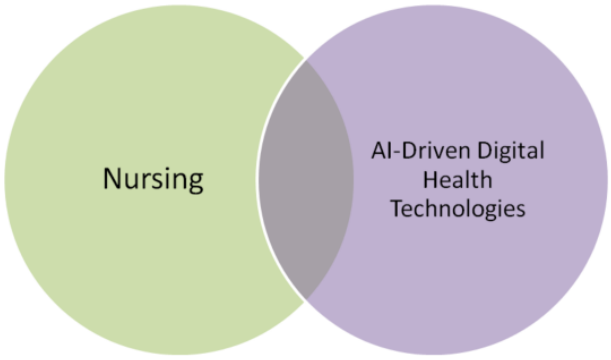
Within the nursing profession, the delivery of compassionate care is a core and valued historical tenet of nursing theory and practice [16-18]; providing safe, compassionate, competent, and ethical care is a core value of nursing practice, as reflected in numerous international nursing practice frameworks [19-21]. Compassionate care has been described as an empathetic response to suffering that involves person-centered care, meaning treating individuals the way they want to be treated [22]. It can be expressed by nurses using silent presence, active listening, firm touch, a caring and respectful attitude, and a kind manner [23]. Compassionate care involves the relationships among health care providers and their patients and can be influenced by a health systems' infrastructure [3]. With compassionate care as a core tenet of the nursing profession, it is important to reflect on the future influence of AI-driven digital technologies on nursing care.

Goals of the Review

It is anticipated that emerging trends in AI-driven digital health technologies will change the nature of nursing roles and practice [2], and in light of technological advancements, nurses will need to ensure the continued delivery of compassionate nursing care [3]. As such, it is important to understand how AI-driven digital health technologies are changing nursing roles and affecting patient and caregiver experiences with nursing care. Nursing students and licensed nurses will need to be equipped to provide care in a technological world, and the influence of AI-driven digital health technologies on nursing education will require examining. Finally, to ensure that AI-driven digital health technologies promote compassionate care (as a core tenet of nursing care) rather than hinder it, it will be important to understand how nurses are involved in the co-design of AI-driven digital health technologies. To explore these focus areas, this scoping review aims to summarize the findings of 4 research questions that explore the relationships between nurses,

patients, and AI-driven digital health technologies (see Figure 1). Furthermore, the findings from this scoping review will be used to reflect on how emerging trends may impact nurses' delivery of compassionate care. The results of this scoping review and the subsequent reflections will be disseminated in a briefing paper and will inform a symposium on the topic of nursing and compassionate care in the age of AI. For this review, a scoping methodology is appropriate due to its exploratory nature and the current literature gap on this research topic. As stated by Tricco et al [24], scoping reviews aim to synthesize evidence and assess the scope of literature on a topic, which is the objective of this review due to the emerging nature of this topic.

Figure 1. Main concepts explored in the review.



Methods

Scoping Review

Standardized reporting guidelines outline items that should be included in research reports to enhance methodological transparency [24]. This protocol was developed using guidance for relevant items from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) [24], as a protocol guidance document for scoping reviews has not been developed. This scoping review is registered in the Open Science Framework database (registration doi: 10.17605/OSF.IO/RTFJN) [25].

This scoping review follows the methodological framework proposed by Arksey and O'Malley [26] and further advanced by Levac et al [27]. This framework delineates 6 steps to map the extent and range of material on a research topic [27], providing clarity on what is known and not known on a topic and situating this within policy and practice contexts [28].

Step 1: Identify the Research Question

A steering committee of experts in digital health technologies was assembled to inform the research questions that would best meet the overarching objective of this review. The steering committee consists of advanced practice nurses, nurse researchers, educators, a patient advocate, registered practical nurses, and registered nurses from various health care sectors. This steering committee is co-chaired by 2 doctorally-prepared nurses who possess independent programs of research in digital health and nursing (RB and TR). Additionally, the Associate Director of the Registered Nurses' Association of Ontario (RNAO) Guideline Development and Evaluation team (MB) and RNAO eHealth Program Manager (RW) were consulted during the development of the research questions.

Levac et al [27] recommend scoping review questions be broad in nature; however, concepts should be clearly articulated to establish an effective search strategy. Using these recommendations, the 4 research questions explored in this review were derived through steering committee discussion and consensus. Due to timeframes and feasibility, the number of research questions was limited to 4:

1. What influences do AI-driven digital health technologies have, or are predicted to have, on the patient/caregiver experience of compassionate care delivered by nurses?
2. What influences do emerging trends in AI-driven digital health technologies have, or are predicted to have, on all domains of nursing practice (ie, administration, clinical care, education, policy, and research)?
3. What influences do emerging trends in AI-driven digital health technologies have, or are predicted to have, on nursing education across all domains?
4. What involvement do nurses have, or are predicted to have, in the co-design of AI-driven digital health technologies?

For the purposes of this scoping review, digital health technologies refer only to those technologies that meaningfully incorporate AI. Literature focused on electronic medical records, telehealth systems, genomics, virtual reality devices, and other technologies that do not actively utilize a discernible or definable form of AI will be excluded. For a technology to be identified as "emerging," the author will have used the word "emerging" or a synonym (eg, new, innovative) to describe the technology, or the co-chairs or steering committee will have identified the technology as emerging based on experience in the field. In this review, a nurse refers to any nurse (eg, registered nurse, registered practical nurse, licensed practical nurse, nurse practitioner) working in any of the 5 nursing domains outlined by the Canadian Nurses Association: administration, education, clinical practice, policy, and research [19]. These 5 nursing domains are also applicable to international nursing settings, as identified in similar international nursing practice frameworks [20,21].

Step 2: Identify Relevant Studies

Peer-Reviewed Literature

The databases MEDLINE, Cumulative Index of Nursing and Allied Health Literature (CINAHL), Embase, PsycINFO, Cochrane Database of Systematic Reviews, Cochrane Central, ERIC, Scopus, Web of Science, and Proquest were searched for peer-reviewed literature using searches developed by an information specialist. The search terms (Multimedia Appendix 1) and inclusion/exclusion criteria (Textbox 1) were developed by examining relevant publications and through consultation with the co-chairs, steering committee, and information specialist. Terms relating to "compassionate care" specifically were not included in the search strategy. Through discussion, it was decided that including "compassionate care" as a search term would narrow the yield of results; as stated by Levac et al [27], scoping reviews are broad in nature as the focus is on summarizing the breadth of the evidence. Therefore, the reviewers felt it was important that the search strategy was broad to encompass a large yield of studies. For the purposes of this review, electronic medical records were also excluded since they are a well-established technology and not an emerging trend. Clinical information from electronic medical records that utilize AI-driven technology (ie, advanced clinical decision support systems) will be included. Due to feasibility, timeframes, and the lack of detailed conference proceedings provided, it was decided that they would also be excluded.

Textbox 1. Inclusion and exclusion criteria.

Inclusion Criteria

- Relates to one of the research questions
- Focuses on AI-driven digital health technology
- Clear connection to nursing (can broadly focus on health care workforce but must be transferable to nursing)
- Published after January 1, 2014 (unless it is a seminal article)
- Peer-reviewed or grey literature, including any research study design (ie, randomized controlled trials, systematic reviews and quasi-experimental, observational and qualitative designs), thesis/dissertations, discussion papers, or white papers
- Printed in English
- Accessible for retrieval

Exclusion Criteria

- Studies/articles not relevant to one of the research questions
- No clear focus on AI-driven digital health technologies
- Focus on electronic medical records
- Specific to other health care profession (eg, physicians only)
- Articles published before January 2014 (unless it is a seminal article)
- Conference proceedings
- Clinical trials
- Not available in English

A test search was run in MEDLINE, and the first 100 articles were reviewed by the reviewers and co-chairs for relevancy. Any additional search terms identified as per relevant articles were added to the search string at this stage. Once consensus was reached among the reviewers and co-chairs regarding relevance of the articles in the test search, it was decided that the full search would be run in the remainder of the databases.

Qualitative and quantitative studies were eligible for inclusion. Due to the rapid increase in use of digital health technologies over the last 10 years, initially it was decided that only literature published after January 1, 2009 would be included. However, due to the substantial yield of articles, it was not feasible to screen all articles considering the project timeframes, and the limit was reduced to the last 5 years upon consultation with the committee co-chairs. Additionally, to balance breadth and feasibility when conducting this scoping review, books/book chapters and conference proceedings were excluded, and only literature published in English was included. Keywords related to the 2 main concepts ("artificial intelligence" and "nursing") were used to search databases (see Multimedia Appendix 1). As already stated, to allow for a broader search, "compassionate care" was not used as a search term in the database search; however, when reflecting on the findings, the reviewers utilized an analytic lens examining the literature's relevance to compassionate care, as this is one of the main focuses of the scoping review. Reviewers referred to the definition of compassionate care outlined in the introduction of this paper when assessing each article's relevance to compassionate care.

Grey Literature

Grey literature was retrieved by searching Proquest, CINAHL, and PsychINFO for English theses and dissertations as well as discussion papers written after January 1, 2014. The same search terms used to search the electronic databases for peer-reviewed literature were used. In addition, targeted website searches of

the following websites will be conducted to retrieve relevant white papers: World Health Organization, National Health Service, Office of the National Coordinator for Health Information Technology, Institute for Research on Healthy Public Policy, Canada Health Infoway, Canadian Association of Schools of Nursing, and Healthcare Information and Management Systems Society. The targeted websites were discussed and agreed upon by the steering committee. Targeted website searches will be performed using Google search strings developed by the information specialist and run by the reviewers (CB and LH).

All peer-reviewed and grey literature results will be downloaded into EndNote X7.8 (Clarivate Analytics, Philadelphia, PA) and imported into the Web-based systematic review software Distiller SR (Evidence Partners, Ottawa, Canada) for review. The expert steering committee will also be asked to identify other potentially relevant peer-reviewed and grey literature materials not identified through prior search strategies (ie, "hand-searched" articles).

Step 3: Study Selection

A screening guide developed by the reviewers (CB and LH) will be used to determine if the inclusion and exclusion criteria have been met (Textbox 1). Feedback was obtained from the co-chairs (RB and TR), eHealth Program Manager and project lead (RW), and Associate Director of the Guideline Development and Evaluation team (MB) while developing the screening guide. The 2 reviewers will independently pilot test the screening guide and review the first 100 abstracts before continuing with screening. Results will be discussed, and revisions to the screening guide will be made as needed. An example of an included article and an excluded article will also be presented to the project team to ensure appropriateness of the articles being included. All titles and abstracts will be independently screened by the 2 reviewers using the screening guide, and the reviewers will meet at the beginning, middle, and final stages of the screening process to discuss challenges related to study selection. First, titles and abstracts will be screened for relevance to "AI-driven digital health technologies" and "nursing" and the general inclusion criteria (Textbox 1). Next, any included full-text articles will be independently reviewed by the 2 reviewers for relevance to determine which of the 4 research questions they address. There will be specific inclusion questions in alignment with each of the research questions to direct the reviewer to which question the article could fall under. Discrepancies in study selection will be resolved by consensus, with final decisions being made by the project lead (RW) if consensus between the 2 reviewers cannot be reached. When screening the full-text studies, reviewers will keep in mind person-centered care principles [29], as these are an important element of nursing practice and compassionate nursing care.

Step 4: Charting the Data

The data charting process will start with studies related to research question 1 and follow in a sequential order (research question 2, then 3, lastly question 4). Draft data charting forms will be developed in Microsoft Excel 2007 (Microsoft Corp, Redmond, WA) for each research question by the 2 reviewers. This form will be reviewed by MB, and the final form will be approved prior to pilot testing. The 2 reviewers (CB and LH) will pilot test the forms by independently charting the data from a representative sample (ie, 5-10 articles per research question) to ensure that consistency is achieved. This sample data charting form will be shared with the co-chairs (RB and TR), and further refinements will be made. Once consistency is achieved and the pilot-tested forms are approved, data from each included full-text article will be charted by one member of the research team and verified by a second member to ensure all relevant data are charted. Levac et al [27] suggest that the development of charting forms is an iterative process and the forms are expected to evolve as literature is reviewed and findings important to the research questions are added to the data fields.

Step 5: Collating, Summarizing, and Reporting the Results

Levac et al [27] recommend that scoping reviews provide a numerical summary of the types of literature retrieved and a descriptive thematic summary of themes arising. Given the expected diverse body of literature, categorical data related to specific elements (ie, study methods, context of study, aim and purpose, key findings) will first be recorded (see Textbox 2); this information will align with the data charting forms for each research question. Study findings will then be synthesized using narrative description. Outcomes will be reported by study type for each research question (ie, qualitative versus quantitative study designs), and themes that emerge for each question will be reported. The findings will also be collated in relation to domains of nursing practice, in order to identify gaps for future research considerations. Pending the results, visual representations of the data may also be created. NVivo 12 (QSR International Pty Ltd, Burlington, MA) and Microsoft Excel 2007 (Microsoft Corp, Redmond, WA) will be used to assist with categorizing and analyzing the data as appropriate.

Textbox 2. Sample data charting elements.

Article Information

- RefID number
- Data charted by (initials)
- Author
- Year
- Study design
- Country
- Aim/purpose

Population

- Nursing designation (registered nurse [RN], registered practical nurse [RPN], licensed practical nurse [LPN], nurse practitioner [NP], student)
- General description of "health care providers"
- Domain of practice or setting

Intervention

- Type of artificial intelligence–driven digital health technology discussed
- Brief description of study

Study Findings

- Key findings related to the research questions
- Relevance to compassionate care

Step 6: Consultation

Levac et al [27] list consultation as a final and mandatory step in the scoping review process. Feedback will be sought from the co-chairs throughout the scoping review process. Once preliminary findings have been identified, a document summarizing the articles included and the themes identified for each research question will be circulated to the steering committee to review. During the consultation, the steering committee members will be asked to reflect on whether the themes identified resonate with their areas of expertise or if there were any themes they expected to see that were not identified. This consultation with the co-chairs and steering committee will occur through video conference call meetings and email. Upon completion of the review, a symposium with key partners and stakeholders will be held to further consult and discuss the findings, foster new delivery models of compassionate care involving digital health technologies, and help build leadership capacity among health care providers.

Results

Electronic database searches were conducted in November 2019, and 14,415 results were retrieved. When the search was limited to the last 5 years, a total of 10,318 articles were retrieved. Title and abstract screening, data charting, and the remaining steps of the scoping review including dissemination (ie, symposium and subsequent briefing paper) aim to be completed by the fall of 2020. Preliminary screening results show that social robots that utilize AI are being used more frequently in long-term care settings with elderly patients [30,31], which may influence the therapeutic relationship between patients and their nurses. Furthermore, clinical predictive models using AI-driven technologies are becoming more advanced and have the potential to positively impact patient care [32,33]; nurses are increasingly becoming involved in the development of these AI-driven digital health technologies [2,12].

Discussion

Preliminary Findings

Digital health technology is frequently conceptualized as being at odds with humanistic, compassionate care, yet both play important roles in the delivery of health care [34]. This dualistic portrayal may prevent health care providers from recognizing the ways in which technology is ubiquitous in their own clinical practice, shaping their relationships with patients in ways that can be difficult to see [34]. The entwinement of technology in health care creates risks and possibilities; health care providers can use technology positively in the provision of patient care, and they can also become dependent on technology to the extent that they lose their humanness [34-36]. Without critical recognition of their entwined human and technology relationships, health care providers are at risk of continuously struggling against technology or being governed by it [34]. However, with conscious recognition of the ways in which technology is enmeshed in clinical practice, health care providers can preserve their humanness, recognize the embodied experiences of patients, and ensure compassionate care is actualized in a technological world [34].

Compassionate care is fundamental to the nursing profession [16-18], and without this core tenet, the nursing profession as it exists now may cease to remain. By critically examining the influence of technology on nursing practice and, more specifically, the influence of AI on compassionate nursing care, the profession can plan for its future trajectory in a technological world.

After collating and analyzing the findings of this scoping review, a briefing paper will be written. The briefing paper will summarize the findings in a narrative fashion, organized in a way that addresses the implications for nursing, health care policy, and future research. A manuscript containing the final analysis will be written and submitted for publication, and the PRISMA-ScR will be completed and submitted with the paper [24].

Limitations

One limitation of this scoping review was the inability to search engineering and computer science databases, due to accessibility issues and organizational licensing restrictions; this limitation may lead to some gaps in the research findings. Future research encompassing engineering and computer science databases on this topic is advised.

In addition, this scoping review did not incorporate the Peer Review of Electronic Search Strategies elements [37]. This process involves peer review of the work of the information specialist, to ensure that the search strategy is appropriate [37]. Typically, this process is only done for systematic reviews, and due to the chosen methodological framework [27] along with timeframes and feasibility, it was not conducted for this review.

Furthermore, the reviewers used percentage agreement when calculating interrater agreement during title and abstract screening, for feasibility purposes, with a percentage agreement rating of 97%. However, it is recognized that this is not as reliable as Cohen's kappa; in future projects, the reviewers will consider using Cohen's kappa for greater interrater reliability.

The lack of quality assessment of the included articles is recognized as another limitation of this scoping review. However, scoping reviews typically do not appraise the quality of evidence in the primary research studies [26]. Furthermore, scoping reviews do not synthesize the research findings based on the relative quality of evidence in favor of a particular intervention; instead, they offer a narrative or descriptive account of the evidence [26]. Although these are potential limitations, scoping reviews can provide a large breadth of research studies in a relatively short amount of time, allowing reviewers to map out the gaps in research as well as summarize and disseminate their findings in a realistic and feasible manner [26]. Finally, given the emerging nature of this topic, performing a quality appraisal may exclude research studies that are more discussion-based or qualitative in nature, thus leading to very limited results.

Conclusions

From a historical perspective, to consider the work of Sandelowski [38], nurses can take one of two viewpoints on technology: technological optimism or technological romanticism. Technological optimism encompasses a positive viewpoint, where one sees the benefits of technology on nursing practice; conversely, nursing romanticism holds a more negative viewpoint, where technology is seen as disruptive and dangerous to nursing practice [38]. In recent years, technology has become increasingly entrenched in nursing practice. Although Sandelowski's viewpoints may not necessarily be true of today's nurses, it is important to recognize both of these perspectives in order to ensure compassionate care is actualized as AI-driven digital health technologies continue to emerge.

Healthcare technology and compassionate care both play vital roles in the provision of nursing care. To our knowledge, this is the first scoping review to examine the influence of emerging trends in AI-driven digital health technologies on all domains of nursing and, more specifically, on the compassionate care that nurses provide. The findings of this scoping review will be relevant for nurse educators, administrators, health care organizations, nursing regulatory bodies, and nursing professional groups preparing nurses for practice in an era of AI. Furthermore, it will provide valuable information on roles of nurses in the co-design of AI-driven digital health technologies. Results of this review will be disseminated in a briefing paper, which will be used to inform a symposium organized by Associated Medical Services Healthcare in partnership with the RNAO.

Acknowledgments

We would like to thank the steering committee for their ongoing contributions to this project: Jocelyn Bennett, Dr. Vanessa Burkoski, Connie Cameron, Roxanne Champagne, Cindy Fajardo, Cathryn Hoy, Vicki Lejambe, Lisa Machado, and Shirley Viaje. We also gratefully thank Marina Englesakis, information specialist at the University Health Network in Toronto, for her assistance in the development of the search strategy. This research is jointly funded by AMS Healthcare and RNAO.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms.
[DOCX File , 719 KB - resprot_v9i4e17490_app1.docx]

References

1. Lexico. Lexico US Dictionary. 2019. Definition of artificial intelligence in English URL: https://www.lexico.com/en/definition/artificial_intelligence [accessed 2019-11-19]
2. Robert N. How artificial intelligence is changing nursing. Nurs Manage 2019 Sep;50(9):30-39. [doi: 10.1097/01.NUMA.0000581404.86147.ac] [Medline: 31460899]
3. Associated Medical Services Healthcare. 2018. Compassion in a technological world: advancing AMS' strategic aims URL: http://www.ams-inc.on.ca/wp-content/uploads/2019/01/Compassion-in-a-Tech-World.pdf [accessed 2019-11-12]
4. Zlotnik A, Alfaro MC, Pérez MCP, Gallardo-Antolín A, Martínez JMM. Building a Decision Support System for Inpatient Admission Prediction With the Manchester Triage System and Administrative Check-in Variables. Comput Inform Nurs 2016 May;34(5):224-230. [doi: 10.1097/CIN.0000000000000230] [Medline: 26974710]
5. Liao P, Hsu P, Chu W, Chu W. Applying artificial intelligence technology to support decision-making in nursing: A case study in Taiwan. Health Informatics J 2015 Jun;21(2):137-148. [doi: 10.1177/1460458213509806] [Medline: 26021669]
6. Abbott MB, Shaw P. Virtual Nursing Avatars: Nurse Roles and Evolving Concepts of Care. Online J Issues Nurs 2016 Aug 15;21(3):7 [FREE Full text] [doi: 10.3912/OJIN.Vol21No03PPT39,05] [Medline: 27857172]
7. Papadopoulos I, Koulouglioti C, Ali S. Views of nurses and other health and social care workers on the use of assistive humanoid and animal-like robots in health and social care: a scoping review. Contemp Nurse 2018;54(4-5):425-442. [doi: 10.1080/10376178.2018.1519374] [Medline: 30200824]
8. Miotto R, Wang F, Wang S, Jiang X, Dudley J. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 2018 Nov 27;19(6):1236-1246 [FREE Full text] [doi: 10.1093/bib/bbx044] [Medline: 28481991]
9. Bresnick J. Health IT Analytics. 2018 Dec 18. Artificial intelligence in healthcare spending to hit $36B URL: https://healthitanalytics.com/news/artificial-intelligence-in-healthcare-spending-to-hit-36b [accessed 2019-11-19]
10. National Health Service: Health Education England. 2019. The Topol review: preparing the healthcare workforce to deliver the digital future URL: https://topol.hee.nhs.uk/ [accessed 2020-01-10]
11. McGrow K. Artificial intelligence: Essentials for nursing. Nursing 2019 Sep;49(9):46-49 [FREE Full text] [doi: 10.1097/01.NURSE.0000577716.57052.8d] [Medline: 31365455]
12. Fritz RL, Dermody G. A nurse-driven method for developing artificial intelligence in smart homes for aging in place. Nurs Outlook 2019;67(2):140-153. [doi: 10.1016/j.outlook.2018.11.004] [Medline: 30551883]
13. McMurray J, Strudwick G, Forchuk C, Morse A, Lachance J, Baskaran A, et al. The Importance of Trust in the Adoption and Use of Intelligent Assistive Technology by Older Adults to Support Aging in Place: Scoping Review Protocol. JMIR Res Protoc 2017 Nov 02;6(11):e218 [FREE Full text] [doi: 10.2196/resprot.8772] [Medline: 29097354]
14. Burkoski V. Nursing Leadership in the Fully Digital Practice Realm. Nurs Leadersh (Tor Ont) 2019 May;32(SP):8-15. [doi: 10.12927/cjnl.2019.25818] [Medline: 31099743]
15. Ihde D. Technology And The Lifeworld: From Garden To Earth (indiana Series In The Philosophy Of Technology). Indianapolis: Indiana University Press; 1990.
16. Gaut D, Leninger M. Caring: The Compassionate Healer. Denver: National League for Nursing Press; 1991.
17. Curtis K. Compassion is an essential component of good nursing care and can be conveyed through the smallest actions. Evid Based Nurs 2015 Jul;18(3):95. [doi: 10.1136/eb-2014-102025] [Medline: 25673277]
18. Chambers C, Ryder E. Compassion And Caring In Nursing. Abingdon, UK: Radcliffe Publishing; 2009.
19. Canadian Nurses Association. Framework for the practice of registered nurses in Canada. 2015. URL: https://tinyurl.com/yxcgu3hn [accessed 2019-11-12]
20. Australian Nursing and Midwifery Federation. Nursing practice. 2018. URL: http://anmf.org.au/documents/policies/P_Nursing_practice.pdf [accessed 2020-01-10]
21. American Nurses Association. Nursing scope and standards of practice (3rd edition). 2015. URL: https://www.iupuc.edu/health-sciences/files/Nursing-ScopeStandards-3E.pdf [accessed 2020-01-10]
22. Sharp S, McAllister M, Broadbent M. The vital blend of clinical competence and compassion: How patients experience person-centred care. Contemp Nurse 2016;52(2-3):300-312. [doi: 10.1080/10376178.2015.1020981] [Medline: 26077823]
23. Perez-Bret E, Altisent R, Rocafort J. Definition of compassion in healthcare: a systematic literature review. Int J Palliat Nurs 2016 Dec;22(12):599-606. [doi: 10.12968/ijpn.2016.22.12.599] [Medline: 27992278]
24. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
25. Buchanan C, Howitt ML, Bamford M. Open Science Framework Registries. 2020. Nursing and Compassionate Care in the Age of Artificial Intelligence: A Scoping Review (Registration) URL: https://doi.org/10.17605/OSF.IO/RTFJN [accessed 2020-01-29]
26. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
27. Levac D, Colquhoun H, O'Brien K. Scoping studies: advancing the methodology. Implement Sci 2010;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]

28. Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. Health Res Policy Syst 2008 Jul 09;6:7 [FREE Full text] [doi: 10.1186/1478-4505-6-7] [Medline: 18613961]
29. Picker Institute Europe. Principles of person centered care. URL: https://tinyurl.com/v9k5mep [accessed 2020-01-30]
30. Moyle W, Bramble M, Jones C, Murfield J. Care staff perceptions of a social robot called Paro and a look-alike Plush Toy: a descriptive qualitative approach. Aging Ment Health 2018 Mar;22(3):330-335. [doi: 10.1080/13607863.2016.1262820] [Medline: 27967207]
31. Kriegel J, Grabner V, Tuttle-Weidinger L, Ehrenmüller I. Socially Assistive Robots (SAR) in In-Patient Care for the Elderly. Stud Health Technol Inform 2019;260:178-185. [Medline: 31118335]
32. Kaewprag P, Newton C, Vermillion B, Hyun S, Huang K, Machiraju R. Predictive models for pressure ulcers from intensive care unit electronic health records using Bayesian networks. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):65 [FREE Full text] [doi: 10.1186/s12911-017-0471-z] [Medline: 28699545]
33. Li H, Lin S, Hwang Y. Using Nursing Information and Data Mining to Explore the Factors That Predict Pressure Injuries for Patients at the End of Life. Comput Inform Nurs 2019 Mar;37(3):133-141. [doi: 10.1097/CIN.0000000000000489] [Medline: 30418245]
34. Lapum J, Fredericks S, Beanlands H, McCay E, Schwind J, Romaniuk D. A cyborg ontology in health care: traversing into the liminal space between technology and person-centred practice. Nurs Philos 2012 Oct;13(4):276-288. [doi: 10.1111/j.1466-769X.2012.00543.x] [Medline: 22950731]
35. Haraway D. A Cyborg Manifesto: Science, Technology, and Socialist Feminism in the Late Twentieth Century. In: Simians, Cyborgs And Women: The Reinvention Of Nature. New York: Free Assn Books; 1991:149-181.
36. Almerud S, Alapack RJ, Fridlund B, Ekebergh M. Beleaguered by technology: care in technologically intense environments. Nurs Philos 2008 Jan;9(1):55-61. [doi: 10.1111/j.1466-769X.2007.00332.x] [Medline: 18154637]
37. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol 2016 Dec;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
38. Sandelowski M. (Ir)reconcilable differences? The debate concerning nursing and technology. Image J Nurs Sch 1997;29(2):169-174. [doi: 10.1111/j.1547-5069.1997.tb01552.x] [Medline: 9212515]

Abbreviations

AI: artificial intelligence.
CINAHL: Cumulative Index of Nursing and Allied Health Literature.
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews.
RNAO: Registered Nurses' Association of Ontario.

Edited by G Eysenbach; submitted 16.12.19; peer-reviewed by J McMurray, G Dermody, C Garritty, G Strudwick; comments to author 09.01.20; revised version received 03.02.20; accepted 21.03.20; published 16.04.20.

Please cite as:
Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M
Nursing in the Age of Artificial Intelligence: Protocol for a Scoping Review
JMIR Res Protoc 2020;9(4):e17490
URL: http://www.researchprotocols.org/2020/4/e17490/
doi:10.2196/17490
PMID:32297873

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org